JOURNAL OF ATMOSPHERIC AND TERRESTRIAL PHYSICS

EDITORIAL ADVISORY BOARD

H. O. G. ALFVÉN–STOCKHOLM · **J. BARTELS**–GÖTTINGEN · **W. J. G. BEYNON**–ABERYSTWYTH · **P. M. S. BLACKETT**–LONDON · **H. G. BOOKER**–ITHACA, N.Y. **E. G. BOWEN**–SYDNEY · **SIR DAVID BRUNT**–LONDON · **R. BUREAU**–BAGNEUX **S. CHAPMAN**–ALASKA · **J. EGEDAL**–COPENHAGEN · **V. F. HESS**–NEW YORK · **E. O. HULBURT**–WASHINGTON, D.C. · **H. ISRAEL**–AACHEN · **D. K. MENZEL**–CAMBRIDGE, MASS. · **S. K. MITRA**–CALCUTTA · **M. NICOLET**–UCCLE · **E. H. NORINDER**–UPSALA · **N. V. PUSHKOV**–MOSCOW · **J. A. RATCLIFFE**–CAMBRIDGE · **O. E. H. RYDBECK**–GOTHENBURG · **J. C. W. SCOTT**–OTTAWA · **SIR GEORGE SIMPSON**–LONDON · **H. SOLBERG**–OSLO · **J. VELDKAMP**–DE BILT · **M. WALDMEIER**–ZÜRICH · **J. ZENNECK**–ALTHEGNENBERG

EDITOR-IN-CHIEF

SIR EDWARD APPLETON, F.R.S.

EDINBURGH

OL.
17
9/60

VOLUME 17

1959–1960

PERGAMON PRESS

LONDON • NEW YORK • PARIS • LOS ANGELES

J. Atmosph. Terr. Phys.

VOL.
17
1959/6

CONTENTS

Nos. 1/2

H. J. A. Chivers and J. S. Greenhow: Auroral ionization and the absorption and scintillation of radio stars 1

H. J. A. Chivers and H. W. Wells: Observations of unusual radiofrequency noise emission and absorption at 80 Mc/s 13

O. Hachenberg and A. Krüger: The correlation of bursts of solar radio emission in the centimetre range with flares and sudden ionospheric disturbances 20

M. D. Vickers: The calculation of the m.u.f. factor for a non-parabolic ionospheric layer 34

L. Wallace: An analysis of a spectrogram of the aurora of 11 February 1958, in the wavelength range 3710–4420 Å 46

C. M. Minnis and G. H. Bazzard: A daily index of solar activity based on *E*-layer ionization (July 1957–December 1958) 57

D. P. Hale: Characteristics of the visual aurorae at Byrd Station, Antarctica, during 1957 65

F. H. Hibberd and J. A. Thomas: The determination of the electron distribution in the upper ionosphere from satellite doppler observations 71

N. C. Gerson and W. H. Gossard: Sweepers 82

A. J. Hirsh: The electron density distribution in the *F*-region of the ionosphere . . 86

J. E. Titheridge: The calculation of real and virtual heights of reflection in the ionosphere 96

J. E. Titheridge: The use of the extraordinary ray in the analysis of ionospheric records 110

J. E. Titheridge: Ionization below the night-time *F*-layer 126

O. C. Jones, R. S. Maddever and J. H. Sanders: Sealed-room experiments on the equilibrium of ionization in air 134

J. A. Lawrie: Rapid fluctuations during magnetic disturbance 145

W. R. Piggott and J. Bhattacharyya: On the mode of propagation in the *E*-layer . 150

F. H. Northover: The propagation of electromagnetic waves in ionized gases (with special reference to "whistlers")—I 158

F. H. Northover: The propagation of electromagnetic waves in ionized gases (with special reference to "whistlers")—II 170

Book Review

Cosmic electrodynamics (Cambridge monographs on mechanics and applied mathematics) by J. W. Dungey (Reviewed by Sir Edward Appleton) 179

Papers to be published in future issues 180

No. 3

H. J. A. Chivers: The simultaneous observation of radio star scintillations on different radio frequencies 181

W. C. Bain: The azimuth distribution of oblique reflections from meteor trails and its relation to meteor radiant distributions 188

H. N. Rundle, D. M. Hunten and J. W. Chamberlain: Resonance scattering by atmospheric sodium—VII. Measurement of the vertical distribution in twilight . 205

M. L. White: Thermal and gravitational atmospheric oscillations—ionospheric dynamo effects included 220

D. M. Schlapp: An attempt to measure the collision frequency of electrons in the *F*-region of the ionosphere 246

Papers to be published in future issues 254

OL.
17
9/60

No. 4

K. C. Yeh and O. G. Villard, Jr.: Fading and attenuation of high-frequency radio waves propagated over long paths crossing the auroral, temperate and equatorial zones 255

B. Chatterjee: A note on fourth reflection condition in the ionosphere . . . 271

L. R. Megill: Photometric observations of the twilight glow [OI] 5577 and [OI] 6300 . 276

A. Onwumechilli: Fluctuations in the geomagnetic horizontal field near the magnetic equator 286

D. M. Hunten: Resonance scattering by atmospheric sodium—VIII. An improved method of deducing the vertical distribution 295

D. H. McPherson and A. Vallance Jones: A study of the latitude dependence of OH rotational temperatures for Canadian stations 302

Y. S. N. Murty and S. R. Khastgir: Group refractive index of the ionosphere at low frequency 309

D. W. G. Chappell: A relationship between the life of M-regions and the rate of change of solar activity 315

J. Veldkamp: A giant geomagnetic pulsation 320

J. G. J. Scholte: On the theory of giant pulsations 325

S. Shanack: A theoretical current density ansatz for the quiet day solar semi-diurnal tidal mode of oscillation of the ionosphere 337

Research Notes

M. S. V. Gopala Rao, B. Ramachandra Rao and P. Ramachandra Rao Pant: Correlation of spread-F activity with F-region height changes 345

M. Gadsden: Colour photography of the aurora 347

Papers to be published in future issues 350

VOL.
17
1959/6

AUTHOR INDEX

(B.R.) denotes a book review

Appleton E. V. (B.R.) 179
Bain W. C. 188
Bazzard G. H. 57
Bhattacharyya J. 150
Chappell D. W. G. 315
Chatterjee B. 271
Chivers H. J. A. 1, 13, 181
Gadsden M. 347
Gerson N. C. 82
Gopala Rao M. S. V. 345
Gossard W. H. 82
Greenhow J. S. 1
Hachenberg O. 20
Hale D. P. 65
Hibberd F. H. 71
Hirch A. J. 86
Hunten D. M. 295
Jones O. C. 134
Khastgir S. R. 309
Krüger A. 20
Lawrie J. A. 145
McPherson D. H. 302
Maddever R. S. 134
Megill L. R. 276
Minnis C. M. 57
Murty Y. S. N. 309
Northover F. H. 158, 170
Onwumechilli A. 286
Piggott W. R. 150
Ramachandra Rao B. 345
Ramachandra Rao Pant P. 345
Rundle H. N. 205
Sanders J. H. 134
Schlapp D. M. 246
Scholte J. G. J. 325
Shanack S. 337
Thomas J. A. 71
Titheridge J. E. 96, 110, 126
Vallance Jones A. 302
Veldkemp J. 320
Vickers M. D. 34
Villard, Jr. O. G. 255
Wallace L. 46
Wells H. W. 13
White M. L. 220
Yeh K. C. 255

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 1 to 12. Pergamon Press Ltd. Printed in Northern Ireland

Auroral ionization and the absorption and scintillation of radio stars

H. J. A. CHIVERS and J. S. GREENHOW
University of Manchester, Jodrell Bank Experimental Station

(*Received* 23 *March* 1959)

Abstract—The absorption of the radiation from extra-terrestrial radio sources, when observed at low elevations to the north of Jodrell Bank (Geomagnetic latitude 56°), is discussed. At lower culmination the Cygnus radio source is at an elevation of only 4°, and the line of sight to the source then passes through the ionosphere near the zone of maximum auroral activity. Combined observations of the source intensity at 36 Mc/s and 100 Mc/s, and of radar back-scatter echoes at 36 Mc/s, have been made using the 250 ft steerable telescope. The absorption and certain back-scatter echoes can be related to an ionized layer at a height of the order of 100 km, associated with auroral activity. It is shown that with a high sensitivity radar equipment at 36 Mc/s, weak auroral-type echoes from irregularities in the *E*-region aligned along the earth's field occur on about 40 per cent of all days at quite low geomagnetic latitudes.

The suggestion is made that the scintillation of radio stars observed from latitudes similar to that of Jodrell Bank is not due entirely to irregularities in the *F*-region, but that there is a large contribution from these aligned irregularities at *E*-region heights.

1. INTRODUCTION

THE absorption of very high frequency radio waves in the ionosphere cannot be investigated by normal pulse-sounding techniques, as the ionized layers are not sufficiently dense to reflect waves in the frequency range from 30 to 100 Mc/s. However, the extra-terrestrial radio sources behave as generators of constant power at these frequencies, enabling the attenuation of the radio waves as they pass once through the complete ionosphere to be measured (CHAPMAN and LITTLE, 1957).

During the period January to February 1958, it was found that at 79 Mc/s the radio star in Cygnus apparently ceased to scintillate near lower culmination, when observed from the latitude of Jodrell Bank (53°14′N). This behaviour was completely at variance with the expected behaviour of the star, as previously the source always showed intense scintillations at lower culmination (LITTLE and MAXWELL, 1951). Thus, for a short period in March 1958 and April 1958, whilst the 250 ft steerable telescope at Jodrell Bank was equipped for radar observations on the artificial earth satellite 1957β (DAVIES *et al.*, 1959) the receiving equipments were used between successive satellite transits to make a series of intensity comparisons between the radio sources in Cygnus (19N4A), and Cassiopeia (23N5A). Since both sources attain high elevation at Jodrell Bank, but Cassiopeia reaches a minimum elevation of 22° and Cygnus passes through lower culmination at only 4° elevation, it was hoped to investigate the variation of ionospheric absorption effects with elevation by making observations on the two sources. When Cygnus is observed at lower culmination from Jodrell Bank, the line of sight passes through the ionosphere near the latitude of the zone of maximum auroral activity and enhanced ionospheric effects can be expected.

Back-scatter observations were also made using the 36 Mc/s radar equipment already connected to the telescope, in order to search for radio echoes from the ionospheric irregularities responsible for the absorption and scintillation effects.

2. Techniques

2.1. *Radio star observations*

Two frequencies (100 Mc/s and 36 Mc/s) were available for the experiment, using crossed dipole feeds in the 250 ft telescope. The half power beamwidths at 36 Mc/s were $\pm 3°$ in elevation and $\pm 4\frac{1}{2}°$ in azimuth, and at 100 Mc/s, $\pm 1\frac{1}{2}°$ in elevation and $\pm 1°$ in azimuth. For noise measurements the rectified outputs of the radar receivers, which had bandwidths of only a few kilocycles, were passed through d.c. amplifiers and presented on strip chart using pen-recording milliammeters.

The observations were made by positioning the telescope at the declination of Cassiopeia but 1 hr in right ascension away from the source. With the pen recorder driving at 1 in./min, the telescope was held in this position for about 3 min to give a base level on the chart. The telescope was then driven in right ascension on to the source where it was fixed for a few minutes before being driven off through 1 hr of R.A. so that the base level could be checked. This routine was carried through on the Cygnus source, and if time allowed, was repeated. Most observations were made in the early evening hours, and the results cover a series of elevation angles from 4° to 50°.

The Cygnus observations obtained with an array of four Yagi aerials at 79 Mc/s, during routine measurements of ionospheric drifts, are also discussed. This system has been described previously (Maxwell and Dagg, 1954).

2.2. *Radar observations*

The pulsed transmitter operated at a frequency of 36 Mc/s with a peak power of 10 kW, and a pulse length up to 300 μsec. The pulse rate was 18·75/sec, giving range ambiguities every 8000 km. At this frequency the gain of the telescope is approximately 300 compared with a half-wave dipole. Intensity modulated range–time and amplitude–time displays were photographed. A range–time base of variable length up to 3000 km was used, and 50 km range markers were presented. As back-scatter echoes were frequently present during the noise measurements, the transmitter was normally turned off while actually recording the source intensities. A section of the film record is illustrated in Fig. 1, and shows an example of the back-scatter echoes from auroral ionization.

3. Results

3.1. *Attenuation of Cygnus at lower culmination on* 36 Mc/s

In order to determine the intensity of a source with a total power equipment, it is necessary to extrapolate the background radiation using the known beamwidth of the aerial. As Cygnus lies on a steeply sloping part of the galactic background this analysis presents some difficulty, and intensity measurements can only be made to an accuracy of approximately 10 per cent at 36 Mc/s although at 100 Mc/s this error is much reduced. Observations on the Cassiopeia source do not suffer from the same limitations.

For each run the estimated mean intensity of Cygnus has been determined as a fraction of the intensity of Cassiopeia. A true ratio of 0·55 has been assumed for no attenuation (Lamden and Lovell, 1956), and this measurement was

VOL.
17
1959/6

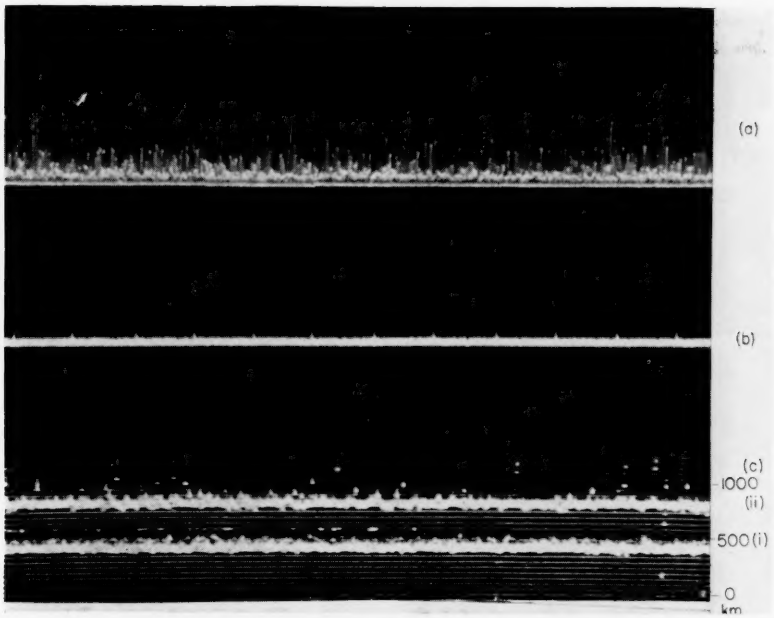


Fig. 1. Section of radar film record. (a) Echo amplitude time display; (b) 1 sec time marks; (c) intensity modulated range–time display showing (i) weak auroral echoes at a range of 330 km, and a (ii) ground scatter echo at twice this range.

OL.
17
9/60

VOL.
17
1959/6

confirmed by a few high elevation observations made during the course of the experiment. The assumption has been made that the mean flux from Cassiopeia is unattenuated since it is always above 22° elevation. This assumption is not quite correct, since even at this elevation some fading has been observed at 36 Mc/s. The measured ratios therefore give the maximum intensity of the Cygnus source in terms of Cassiopeia.

The curve showing the intensity ratio of Cygnus to Cassiopeia at 36 Mc/s, as a function of time from transit of Cygnus, is given in Fig. 2(a). It is clear that the Cygnus source suffered appreciable attenuation (about 7·5 dB) when near lower culmination and did not reach full intensity until about 6 hr from lower culmination when it is at an elevation of approximately 30°. There is a considerable day to day variation in the intensity at a given time, as well as large short period variations with periods of a few minutes to an hour or so. On occasions the source was only attenuated by 2-3 dB near lower culmination, while at other times it was almost undetectable. The short-term variation in signal strength is shown in Fig. 3 which was obtained by following the Cygnus source continuously for about 1 hr. The rapid increase in signal at 2300 hours occurred as the telescope was driven onto the source position. Variations of from 2 to 1 in intensity are observed during this interval, when the mean elevation of the source was approximately 14°. The scintillation amplitude varies with the mean intensity, showing that the depth of scintillation remains roughly a constant fraction of source intensity.

OL.
17
9/60

3.2. 100 Mc/s *observations*

Fig. 2(b) shows the mean intensity ratio of Cygnus to Cassiopeia at 100 Mc/s as a function of time from transit. The greatest mean attenuation (of about 1 dB) is seen to occur when the source is at lower culmination, although the observed intensity at this frequency also showed very large short period variations. These variations were distinctly slower than the normal scintillations which were present at the same time (Fig. 4(a)). Similar fading, with two distinct periods clearly superimposed has been observed when observing Cassiopeia on 36 Mc/s at an elevation as high as 23°. An example is shown in Fig. 4(b) where the mean intensity is seen to fall below half the normal value for a period of about 2 min.

3.3. 79 Mc/s *observations of Cygnus scintillations*

During routine measurements of ionospheric drifts by the radio star scintillation spaced receiver method (MAXWELL and DAGG, 1954) using a radio frequency of 79 Mc/s, it was noticed that when Cygnus was observed near lower culmination the amplitude of the fluctuations decreased noticeably within about 1 hr each side of lower culmination. Because of time constants in the receiver the apparatus was not capable of observing changes in flux density which took longer than 10 min, so that gradual absorption of the source would not be noticed. However, drift observations using Cygnus near lower culmination have been made on thirty-one occasions between January 1958 and May 1958 and on only six occasions were fluctuations with a mean amplitude greater than 20 per cent of the estimated unattenuated source intensity observed. These results are contrary to those of

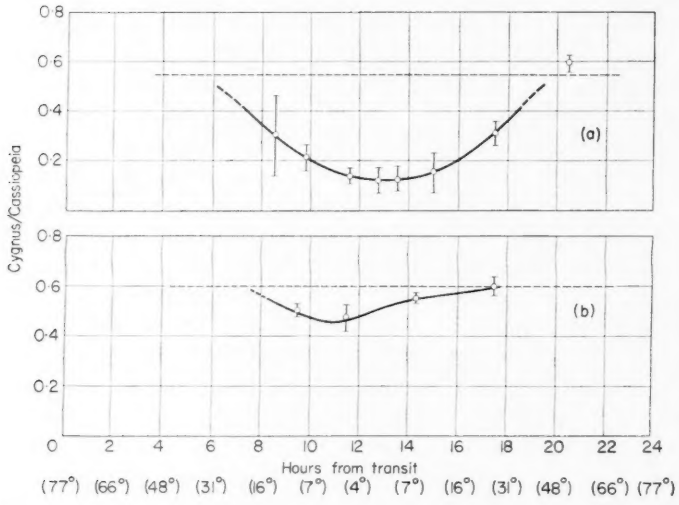


Fig. 2. The intensity ratio of Cygnus to Cassiopeia as a function of hours from upper transit: (a) 36 Mc/s; (b) 100 Mc/s. The figures in brackets indicate the source elevation. The vertical lines show the probable errors in each point.

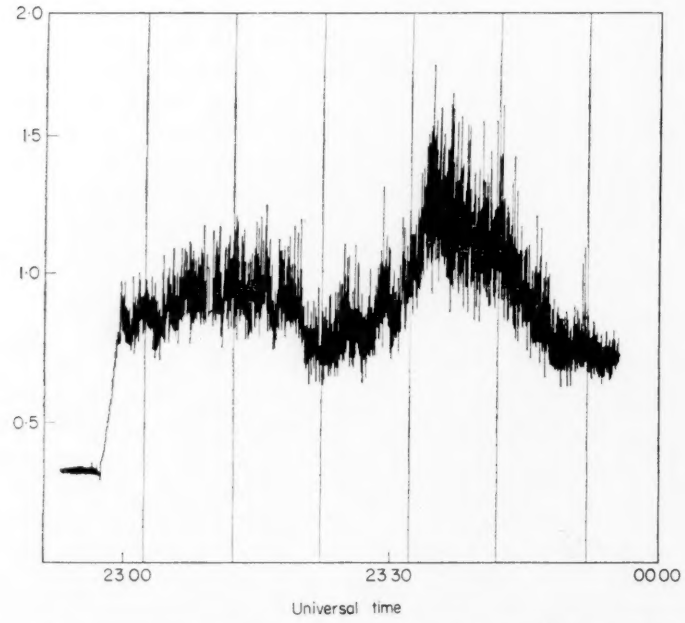
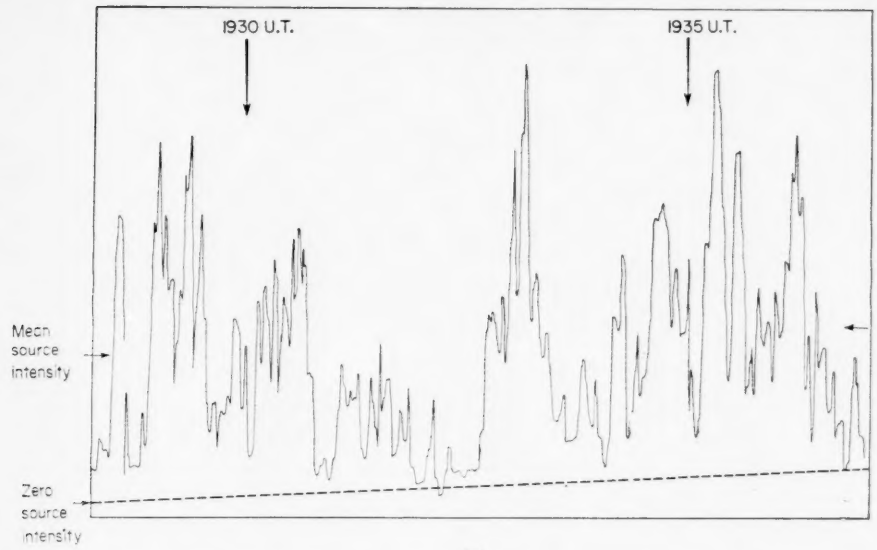
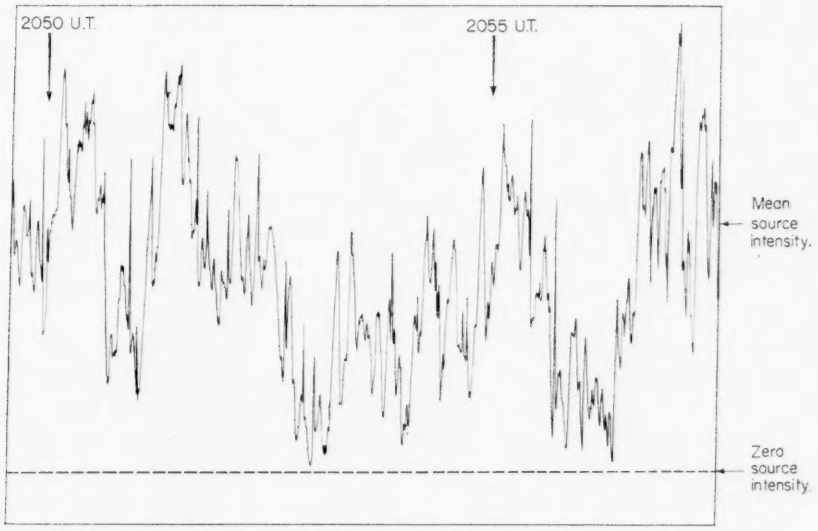


Fig. 3. 36 Mc/s record of Cygnus showing variability in source intensity on an arbitrary scale. The rapid increase in intensity to the left of the record occurred as the telescope was driven onto the source position.



(a)



(b)

Fig. 4. Records showing rapid variation in source intensity. The dashed line indicates zero source intensity on an arbitrary scale. (a) Cygnus, 100 Mc/s, 11 April 1958. Source elevation 5°; (b) Cassiopeia, 36 Mc/s, 11 April 1958. Source elevation 23°.

previous workers (LITTLE and MAXWELL, 1951) who obtained a marked increase in scintillation amplitude when Cygnus was within 2 hr of lower culmination. Also the two wavelength observations described in the previous paragraphs show that near lower culmination, scintillations with a mean depth equal to the observed source intensity are invariably obtained regardless of the attenuation occurring

at the time. It may be assumed, therefore, that the mean fluctuation amplitude of the Cygnus source near lower culmination is in fact large and that the apparent reduction is due to a decrease in the incident flux from the source rather than a decrease in scintillation amplitude. Under these conditions a measure of the scintillation amplitude becomes a measure of source intensity, and this has been plotted on an arbitrary scale in Fig. 5 as a function of time after transit. The results have been divided into two groups: (a) January–March, and (b) April–May.

Once again there is a marked attenuation near lower culmination, the amount of attenuation varying considerably between the two series of observations. From January to March the average attenuation is 6 dB, falling to only 1·8 dB in April–May.

3.4. *Variation of attenuation with frequency*

The 79 Mc/s observations for April–May may be combined with the 36 Mc/s and 100 Mc/s measurements obtained using the 250 ft telescope to investigate the frequency dependence of the Cygnus attenuation at lower culmination. The attenuations at 36, 79 and 100 Mc/s are, respectively, 7·5, 1·8 and 1·0 dB, which within experimental errors are in the ratios required for simple non-deviative absorption where the absorption coefficient varies inversely as the square of the frequency. The possible nature and location of the absorbing region will be discussed later in the paper.

VOL.
17
1959/6

3.5. *Radar observations*

At intervals between the radio-star scans, observations were made with the 36 Mc/s radar apparatus. The sensitivity of this equipment is such that at a range of 300 km an ionospheric irregularity with a scattering cross-section of only 0·1 m^2 would give a detectable echo. Azimuth and elevation scans were carried out in the northern quadrants, and additional observations were made when the telescope was changing the position between the two radio sources. Several types of echo were observed, of which the most frequent indicated the presence of auroral-type ionization.

3.5.1. *Auroral ionization aligned along the earth's magnetic field.* This consists of irregularities in the density of ionization which are aligned along the earth's local magnetic field, and radar echoes can only be obtained from them when the line of sight is very nearly at right angles to the field lines (Harrison and Watkins, 1958). Because of this specular property, echoes at any particular azimuth, will only be obtained at a specific elevation. Hence, echoes from this type of ionization can be easily identified, using a narrow-beam aerial. For ionization at a height of 115 km, it is only possible to satisfy the condition of specularity at some elevation below 20° between azimuths 062° west of north to 040° east of north when observing from Jodrell Bank.

The auroral echoes observed with normal survey equipments are usually recorded during periods of enhanced solar and magnetic activity, and are often concurrent with visual aurora. However, when observations are made using the telescope at 36 Mc/s, weak auroral echoes are observed very frequently. During the radio-star absorption measurements, echoes were obtained at some time on

approximately 40 per cent of all days when observations were made, usually between 1800 and 0600 hours local time. These echoes were very weak, with corresponding scattering cross-sections as low as 0·5 m^2, but they indicate the presence of an almost continuous influx of ionizing particles of the type associated with visual aurora. No unusual magnetic activity was recorded during these displays of auroral echoes (K-index $\leqslant 3$).

3.5.2. *Layer type auroral ionization.* During times when direct back-scatter echoes were obtained from ionization aligned along the earth's field, a second type of

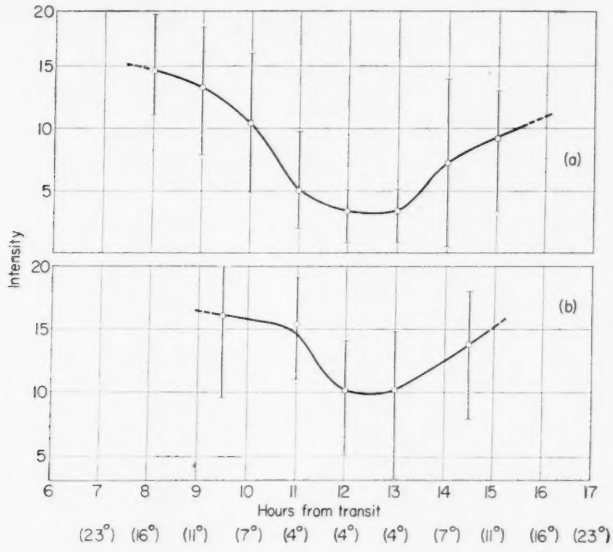


Fig. 5. Mean intensity of Cygnus on an arbitrary scale at 79 Mc/s as a function of time after transit. (a) January 1958–March 1958; (b) April 1958–May 1958. The figures in brackets indicate the elevation of the source. The vertical lines show the spread in the results and include 50 per cent of the individual measurements.

auroral echo was also observed quite frequently. These echoes have been interpreted as scattering from the ground via a horizontal layer at a height of about 100 km, closely associated with the field aligned irregularities (Greenhow, Neufeld and Watkins, unpublished). The close spatial relationship between the two types of ionization is shown by the almost 2 to 1 ratio in ranges between the ground and direct back-scatter echoes (Fig. 1). In order to account for the observed strength of the ground echo, a totally reflecting ionized layer with an equivalent vertical critical frequency of the order of 12 Mc/s must be postulated. The small spread in range of the echoes shows that the horizontal extent of these patches of ionization is of the order of 100 km. Ionization of this type is of particular interest in the present investigation since it would give rise to appreciable absorption at high radio frequencies.

3.5.3. F-*region echoes.* Because of the interest in the location of the ionization causing the scintillation and absorption of radio stars it is important to differentiate between E-region and F-region back-scatter echoes. A well understood

and easily identifiable type of echoes is reflection from the ground via the normal F-layer. At low elevations these echoes occur at ranges of 3000 km and on occasions echoes from up to the fourth hop at 12,000 km have been identified. At 36 Mc/s such echoes normally disappear at night when the F-region electron density falls. No echoes positively identifiable with direct reflections from F-region heights have been obtained, and at 36 Mc/s the most frequent type of night-time echo is that associated with reflection via auroral ionization at E-region heights.

4. Discussion

4.1. *Causes of attenuation in source intensity*

Variations in the apparent intensity of a source could be caused by the following effects:

(a) Absorption of radio waves due to collision processes;

(b) Ionospheric refraction, causing the position of a radio star to deviate from the maximum of the aerial beam;

(c) Divergence of the incident waves at low elevations causing the flux density at the surface of the earth, to be reduced (Bremmer, 1949).

Divergence can be a serious effect at low angles of elevation, when the electron density approaches that required for total reflection. However, 6 dB attenuation at 79 Mc/s at an elevation of 4° would require an F-region equivalent vertical critical frequency of 30 Mc/s (Kerr and Shain, 1951). Similarly, an attenuation of 6 dB at an elevation of 15° at 36 Mc/s would require a critical frequency of approximately 20 Mc/s. Such high critical frequencies are improbable and can in any case be excluded because of the absence of F-region ground scatter echoes at all elevations during the times of observation. A reasonable value of the critical frequency would be 10 Mc/s, and this would give only 1 dB attenuation at 36 Mc/s at an elevation of 4°. This attenuation decreases rapidly with increasing elevation and frequency, and is negligible at 100 Mc/s for all elevation angles.

At frequencies some way from the critical frequency the angle of refraction is small, and in any case refraction could not influence the 79 Mc/s observation since the beamwidth of the aerial was large. As an example, on 36 Mc/s at 4° elevation the refraction would be only 1° for a 10 Mc/s critical frequency. Thus non-deviative absorption is the most probable cause of the observed attenuation, and this affect will now be discussed in greater detail.

4.2. *Non-deviative absorption*

In order to obtain appreciable absorption with reasonable electron densities at frequencies of from 36 to 100 Mc/s, it is necessary to postulate an ionized layer low in the ionosphere. The observed attenuations at an elevation of 4° in April were 7·5, 1·8 and 1 dB at 38, 79 and 100 Mc/s, respectively. At E-region heights the equivalent attenuations for a vertical path would be approximately 2, 0·5 and 0·25 dB. As the line of sight to Cygnus at lower culmination passes close to the auroral zone at ionospheric heights, auroral ionization is the obvious effect with which to associate this absorption. However, the normal height of occurrence of the weak v.h.f. auroral echoes is approximately 115 km, and the total ionization

VOL.
17
1959/6

in the field aligned irregularities is insufficient to cause appreciable absorption at this height.

In order to obtain 0·5 dB attenuation at 79 Mc/s a homogeneous layer 10 km thick would be required with an electron density of about 10^6 electrons/cm^3 centred about 100 km. Evidence for the existence of such dense ionization, associated with auroral activity, is given by the total reflection of 36 Mc/s radio waves which gives rise to the ground back-scatter echoes. An equivalent vertical critical frequency of 12 Mc/s corresponds to an electron density of 2×10^6 cm^{-3}, and this is sufficient to account for the observed absorption for a layer at a height of 100 km. At lower heights correspondingly smaller electron densities would be required.

The small horizontal extent of the patches of ionization, and the intermittent nature of auroral ionization, also readily explains the short period, long-period and day-to-day variations in the amount of absorption.

4.3. *Origin of the scintillations*

4.3.1. *Nature of the diffracting region.* There has been much discussion as to the height at which the irregularities in electron density responsible for radio star scintillations occur. Attempts have been made to associate the irregularities with spread-*F* echoes at a height of 400 km (Dagg, 1957b). However, the heights of the irregularities giving rise to spread-*F* echoes are as yet uncertain, as the electron-density variations could be anywhere below the virtual heights indicated by *h'f* records. Indeed observations of the radio transmissions from artificial satellites show that scintillations are observed when the satellite is as low as 220 km (Kent, 1958). Even this observation only sets an upper limit to the height of the diffracting screen, as scintillations will not be observed when the radio source is as low as the irregular region (Little *et al.* 1956).

Some of the conditions to be satisfied by the ionization responsible for radio-star scintillations are:

(a) The irregularities should be aligned along the earth's magnetic field;
(b) The scintillations should nearly always occur when the line of sight to the star through the irregularities passes close to the auroral maximum zone, irrespective of time of day (e.g. at low elevations to the north of Jodrell Bank);
(c) In lower latitudes the scintillations should occur most frequently at night (e.g. high elevation observations at Jodrell Bank);
(d) The scintillations should correlate with magnetic activity (Dagg, 1957a).

These requirements describe very closely the weak auroral ionization which gives rise to the radio echoes described in Section 3.5.1, and it is tempting to associate these irregularities with scintillation effects. Thus when back-scatter echoes are observed there is positive evidence of electron-density irregularities capable of reflecting appreciable amounts of energy at 36 Mc/s from heights of only 110-120 km. It would be expected, therefore, that when auroral-type echoes are present that the ionization at this level would have a predominant effect on radio-star scintillations. Compared with non-echoing regions the electron densities

will be so great that instead of obtaining the weak scattering normally postulated to explain the scintillations, both the phase and amplitude of the incoming wave will be modified profoundly. When this strong scattering occurs a change in the nature of the scintillations would be expected, and the continuous records of Cassiopeia scintillations made at Jodrell Bank have therefore been examined during the periods of radio-echo observations.

4.3.2. *Correlation of scintillations with weak auroral-echo activity.* Although combined scintillation and radio-echo observations have been made only on about 20 days, a one to one correlation has been found between the type of scintillations and the presence or otherwise of weak auroral echoes. During "no echo" periods the scintillations are either absent or of long period. An example of this type of record is shown in Fig. 6 (a).

The fluctuation rate is only about 1/min, and the intensity variations are given by a displaced Gaussian or Rayleigh distribution (Chivers, 1959). When echoes are present the whole nature of the scintillations change as shown in Fig. 6 (b). The fluctuation rate increases many times, so much so that the individual scintillations cannot be resolved. When examined on a high-speed recorder the amplitude fluctuations are found to be very different from the "no echo" case. Frequent large peaks and periods of very low intensity are observed, characteristic of the focusing and defocusing effects which would be associated with strong scattering. Scintillations of this type are nearly always recorded when the Cygnus source is observed near lower culmination, when the line of sight passes through the 115 km level at a latitude ($\sim$65°N) where a 100 per cent occurrence of weak auroral ionization would be expected.

It is therefore reasonable to suppose that when field-aligned electron-density irregularities capable of giving radar echoes at high frequencies are present, that these are the irregularities primarily responsible for the fast scintillations which are observed. As echoes are observed at some time on about 40 per cent of all days at a geomagnetic latitude of 56°N, this would imply that the scintillation level is often as low as 115 km. At latitudes above 65°N, the scintillation region would almost always be at or below this height. As the height of auroral ionization appears to depend upon the intensity of the aurora it is possible that the more frequent slow fluctuations at middle latitudes are due to very weak ionization above 200 km. Observations of auroral-type echoes from these heights at frequencies of 6 and 17 Mc/s have been described by Peterson *et al.* (1955), although none have been observed with the present 36 Mc/s equipment.

4.3.3. *Drifts determined from discrete auroral echoes and from scintillation observations.* It is interesting to note that the ionospheric drifts deduced from scintillation measurements (Maxwell and Dagg, 1954) are very like those determined from the movements of auroral echoes (Kaiser and Bullough, 1955). Some simultaneous measurements during auroral echo displays have been made (Chivers and Watkins, unpublished), and the drifts determined by both techniques are the same. The similarity of the movements determined by the two methods is thus added confirmation that the mechanism giving rise to auroral echoes is the same as that responsible for the concurrent radio-star scintillations, both occurring at heights of approximately 115 km.

VOL
17
1959/

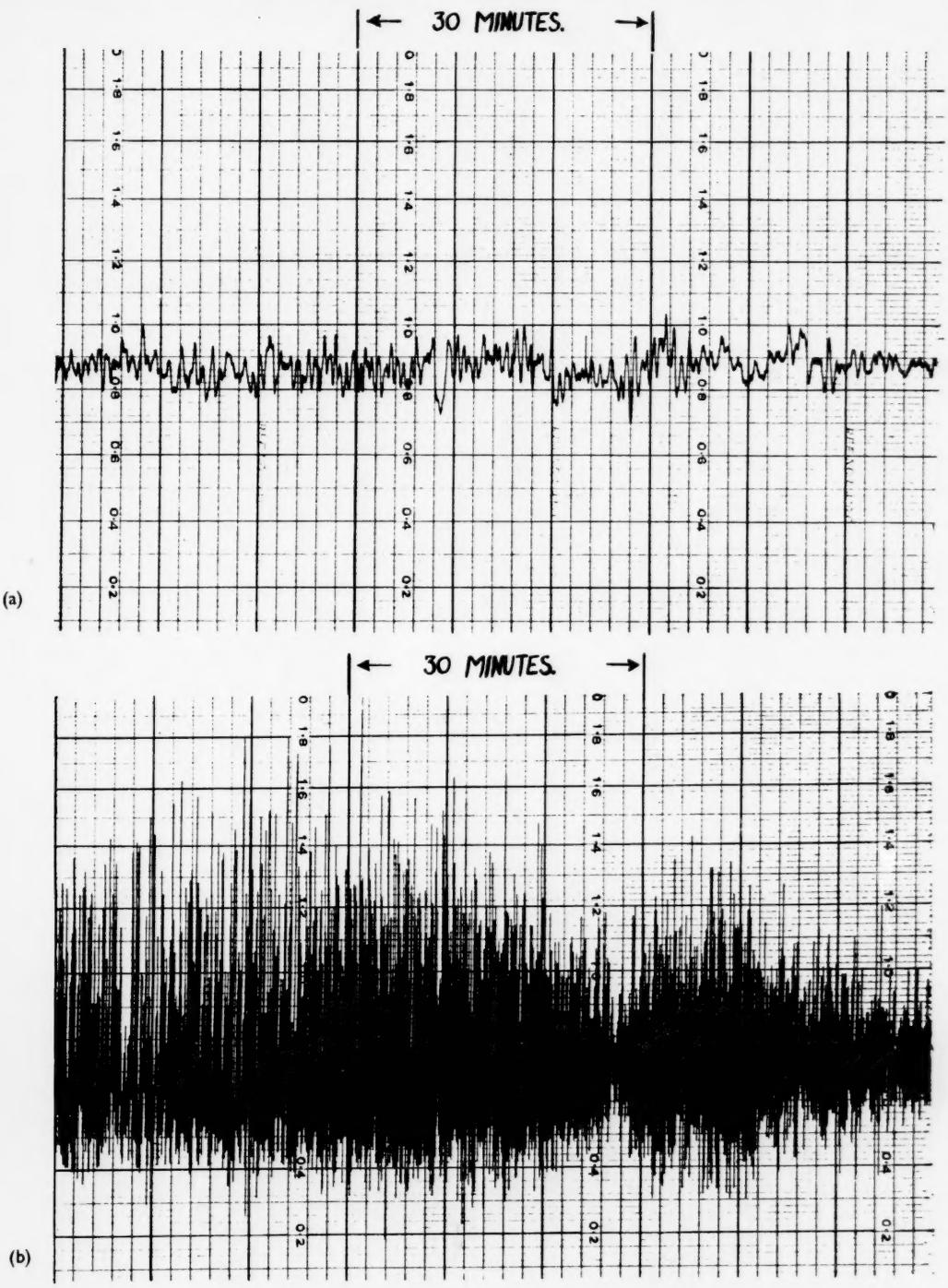


Fig. 6. Records of Cassiopeia scintillations at an elevation of approximately 60°. (a) When no back-scatter echoes are observed; (b) the type of scintillations associated with weak echoes from ionization aligned along the earth's magnetic field at a height of 115 km.

Acknowledgements—The work described in this paper was carried out at the University of Manchester, Jodrell Bank Experimental Station. The writers wish to thank Professor A. C. B. LOVELL, Director of the Station, for his interest in the investigation. They also wish to thank Dr. J. V. EVANS for the use of his 100 Mc/s receiver, and Dr. E. L. NEUFELD who helped in the 36 Mc/s observations.

REFERENCES

BREMMER H.	1949	*Terrestrial Radio Waves.* Elsevier, Amsterdam.
CHAPMAN S. and LITTLE C. G.	1957	*J. Atmosph. Terr. Phys.* **10,** 20.
CHIVERS H. J. A.	1959	*J. Atmosph. Terr. Phys.* In press.
DAGG M.	1957a	*J. Atmosph. Terr. Phys.* **10,** 194.
DAGG M.	1957b	*J. Atmosph. Terr. Phys.* **10,** 204.
DAVIES J. G. *et al.*	1959	*Proc. Roy. Soc.* A **250,** 367.
HARRISON D. P. and WATKINS C. D.	1958	*Nature, Lond.* **181,** 43.
KAISER T. R. and BULLOUGH K.	1955	*J. Atmosph. Terr. Phys.* **6,** 198.
KENT G.	1958	*J. Atmosph. Terr. Phys.* **16,** 10.
KERR F. J. and SHAIN C. A.	1951	*Proc. Inst. Radio Engrs., N.Y.* **39,** 230.
LAMDEN R. and LOVELL A. C. B.	1956	*Phil. Mag.* Ser. 8, **1,** 725.
LITTLE C. G. and MAXWELL A.	1951	*Phil. Mag.* **42,** 267.
LITTLE C. G., RAYTON W. M. and ROOF R. B.	1956	*Proc. Inst. Radio Engrs., N.Y.* **44,** 992.
MAXWELL A. and DAGG M.	1954	*Phil. Mag.* **45,** 551.
PETERSON A. M. *et al.*	1955	*J. Geophys. Res.* **60,** 497.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 13 to 19. Pergamon Press Ltd. Printed in Northern Ireland

Observations of unusual radiofrequency noise emission and absorption at 80 Mc/s

H. J. A. Chivers and H. W. Wells*
Jodrell Bank Experimental Station, University of Manchester

(*Received* 24 *March* 1959)

Abstract—Unusual radiofrequency noise emissions at 80 Mc/s have been identified during periods of solar activity. The noise enhancements may be classified as (1) smooth, bay-like disturbances lasting for approximately 1 hr which occur in both day and night hours, and (2) abrupt increases, often of large but fluctuating amplitude which occur within a few hours of local midnight. The smooth enhancements occur almost simultaneously with the absorption of radiation in a sector of the northern sky. These effects could be caused by transit of high velocity streams of charged particles which produce emission from *F*-region levels and absorption in the *E*-region or below.

The abrupt noise bursts at night are from the northern sky and seem to be associated with pronounced changes in the horizontal component of the earth's magnetic field. The noise may be a form of "auroral" radiation or may arise from propagation of solar noise outbursts from the sunlit to the dark hemisphere.

1. Introduction

In October 1958 a special programme was initiated to study the nature of unusual emissions of noise at radiofrequencies. Interest in such events was stimulated by the independent observation earlier in 1958 of noise enhancements on 80 Mc/s at the Jodrell Bank Experimental Station, University of Manchester and on 87 Mc/s at the Derwood Experimental Station, Carnegie Institution of Washington.

Three independent receiving equipments which were normally used to measure ionospheric drifts by the radio-star scintillation technique were orientated to monitor different sectors of the sky. These were (a) 30° above the southern horizon (b) the zenith and (c) 30° above the northern horizon. A fourth equipment with an aerial driven continuously to follow Cassiopeia 23N5A, provided data at varying directions towards the north. The receivers were operated at staggered frequencies to discriminate against narrow band interference. The maximum frequency separation used was about 0·5 Mc/s which was several times larger than the half-power bandwidths of the receivers.

Continuous operation of the equipments from October 1958 until December 1958 yielded no noteworthy events apart from bursts of solar noise on the south-looking equipment. However, during the interval 3–10 January 1959 a number of distinct examples of both emission and absorption were recorded by the four equipments. This period coincided with the appearance and subsequent development of an unusually large sunspot which gave rise to many solar flares, noise storms, magnetic disturbances and aurorae.

2. Observations

The unusual events observed were of two distinct types and these have been classified as (1) smooth and (2) abrupt disturbances.

* Carnegie Institution of Washington, Washington D.C. (at present at the Jodrell Bank Experimental Station).

2.1 *Smooth disturbances*

These appeared as bulges in the recorded noise level and persisted for 1 hr or more. Both increases and decreases were observed indicating the presence of emission and absorption respectively. It was remarkable that emission in one sector was usually accompanied by absorption in another. The magnitude range of the effects discussed were from 10 to 90 per cent of the normal diurnal change in the galactic background as recorded by the instruments. The strongest events of

Table 1. Smooth disturbances

Date 1959	Aerial				Remarks
	Continuous following	North	South	Zenith	
3 Jan.	—	—	Inc. strong 0530–0645	—	—
4 Jan.	Dec. strong 0840–0950	Inc. medium 0920–0950	—	Inc. strong 0925–1050	Probable magnetic effect in Z
6 Jan.	—	Inc. medium 1645–1830	—	—	—
7 Jan.	Dec. strong 1115–1330	Dec. strong 1115–1245	—	Inc. strong 1045–1230	Max. absorption at different times
7 Jan.	Dec. weak 1845–2000	Dec. weak 1845–2000	Inc. weak 1900–2000	—	Abrupt change in H at 1900; solar flares
9 Jan.	Dec. weak 1530–1645	Dec. medium 1530–1645	Inc. weak 1530–1630	—	Many subflares

this type were recorded during daylight hours and the northern sectors of the sky were the most active both in absorption and emission. The zenith exhibited emission on several occasions, whilst the south showed only a single weak case of emission of this type. During the period 0900–1500 UT daily, the sun was within the beam of the southern aerial and many bursts and noise storms were recorded. These have been excluded from this report. No examples of absorption were recorded in the zenith or south.

Table 1 lists all the events of the above classification. Three are reproduced in Fig. 1(a), (b) and (c), and in each case, the records from all four equipments are shown. The dashed lines indicate the normal levels.

In Fig. 1(a) is shown an event which started at 0840 on 4 January with absorption on the equipment continuously following Cassiopeia. The event lasted nearly 2 hr and at the minimum, the decrease amounted to 50 per cent of the total galactic signal. At this time the aerial was directed 45°E of N at 42° elevation. On

the northern instrument an enhancement of about 15 per cent of the normal change in galactic signal level occurred with a maximum which coincided with the peak absorption on the continuously following (C/F) channel. The southerly equipment indicated normal activity, but the zenith revealed a strong enhancement centred around a later time with an intensity of about 75 per cent of the diurnal galactic range.

Fig. 1(b) features absorption in the C/F (directed at 55°E of N at 56° elevation), absorption in the north, a normal record in the south and emission from the zenith.

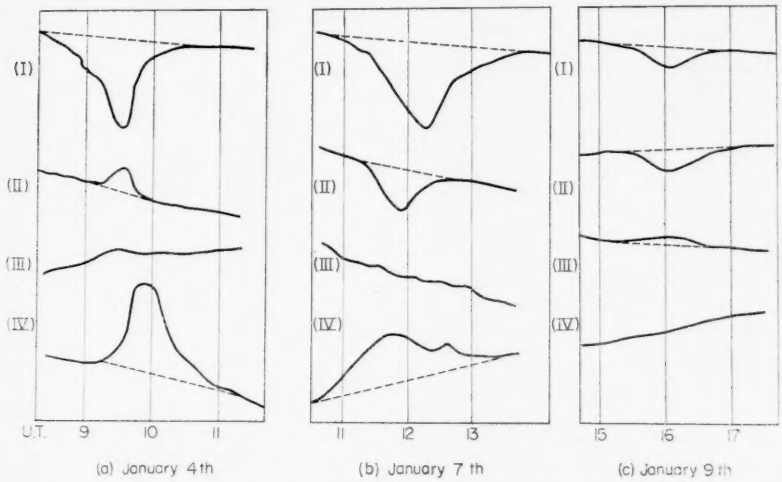


Fig. 1. Examples of smooth disturbances: (i) continuously-following; (ii) north; (iii) south; (iv) zenith.

It is noteworthy that the maxima of the northerly absorption and the zenithal emission occurred at the same time but that the C/F absorption peak did not occur until 25 min later.

The records in Fig. 1(c) show moderate absorption in both northerly sectors with weak emission in the south. The zenithal trace shows no abnormality.

2.2. *Abrupt disturbances*

The abrupt type of noise bursts observed during this period are listed in Table 2. Facsimiles of records from the C/F and northern instruments are presented in Fig. 2(a) and (b). Since the records from the south and zenith were normal at the time they are not reproduced. The third record of each group is from the magnetic horizontal (H) component recorded at Jodrell Bank. The surges in H coincide with the noise bursts and are significant features in the records. The amplitudes of the bursts are from 100 per cent to at least 200 per cent of the diurnal galactic range and were off-scale in some cases. The solid blocks in Fig. 2(a) shown below the magnetogram indicate the times of reported major radio noise bursts on 18 Mc/s and 200 Mc/s.

During the interval discussed, the Jodrell Bank radar equipment recorded only one aurora. This was at 1710 UT on 9 January and was accompanied by an abrupt

but weak enhancement in the north radiometer as listed in Table 2. Verbal reports of a visual aurora later that evening were received but no significant effects were shown on the instruments.

Table 2. Abrupt disturbances

Date 1959	Aerial				Remarks
	Continuous following	North	South	Zenith	
5 Jan.	Inc. strong 2145–2155	Inc. strong 2125–2150	—	—	Flares preceding magnetic effect in H; solar noise storm
8 Jan.	—	Inc. strong 2245–2200	Inc. weak 2245–2255	Inc. strong 2200–2310	Magnetic effect in H
9 Jan.	—	Inc. weak 1700–1715	—	—	Radar aurora, weak
10 Jan.	Inc. strong 2200–2240	Inc. strong 2200–2240	—	—	Subflares
27 Jan.	Inc. strong 0130–0230	Dec. weak 0200–0230	—	—	Class 1 flare, 0038–0056

VOL. 17 1959/6

3. Discussion

In August and September 1958 the writers independently recorded unusual noise enhancements during the night hours which were associated with periods of solar magnetic or auroral activity. These observations were made at the Derwood Experimental Laboratory, Maryland, U.S.A. on 87 Mc/s, and at Jodrell Bank using the continuously-following equipment on 80 Mc/s. It is hoped to deal with these events in a later publication.

A casual survey of the literature reveals several reports of noise bursts during periods which are often associated with aurora. Appleton and Hey (1946) identified a strong noise source in the north with two independent equipments on about 150 Mc/s coinciding with the presence of a magnetic storm. Covington (1947) reported 10 cm noise bursts coinciding with steep changes in the magnetic field. Later, Hartz and his associates (1956) detected noise bursts in the frequency range 32–53 Mc/s which were established as auroral in origin.

Others have searched for such emissions, principally at the higher frequencies, without success. The possibility of the generation of Čerenkov radiation by the passage of high-velocity particles through the normal F-region of the ionosphere has been examined by Ellis (1957) who concludes that the electromagnetic energy would be radiated at frequencies below one megacycle per second. However,

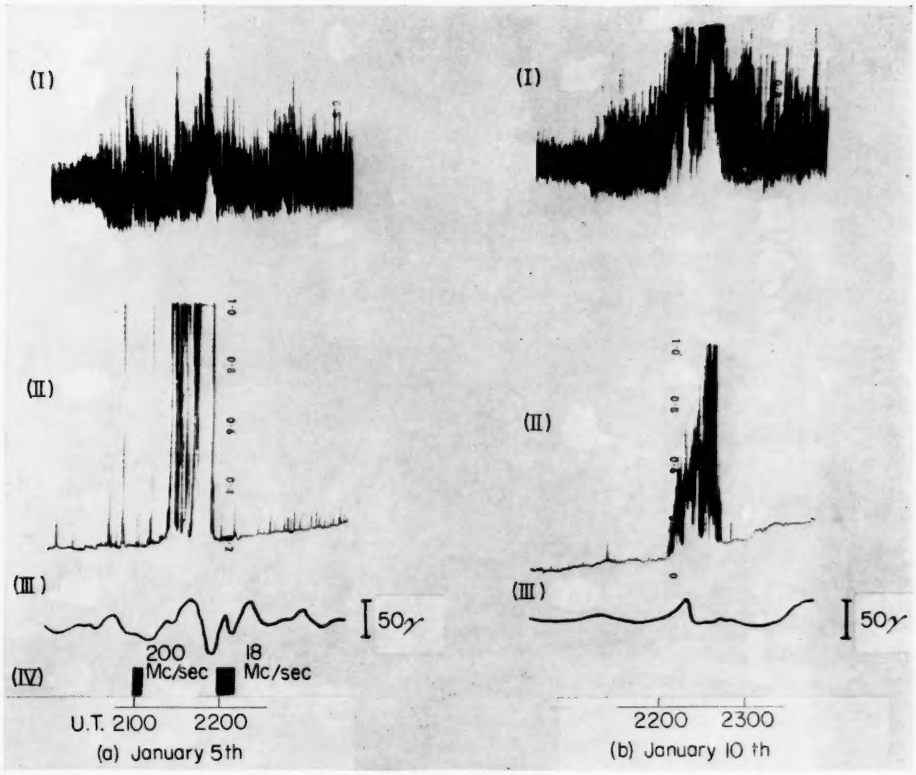


Fig. 2. Examples of abrupt disturbances: (i) continuously-following record with scintillations; (ii) North; (iii) magnetic H-component; (iv) duration of solar noise bursts on sunlit hemisphere at 200 Mc/s and 18 Mc/s.

VOL.
17
1959/6

HARTZ (1958) considers that 500 Mc/s radiation can be generated in this way by the adoption of more generous assumptions.

The absorption phenomena, on the other hand, are in accordance with the well established events which occur very commonly at frequencies below 40 Mc/s notably at high latitudes (CHAPMAN and LITTLE, 1957; CHIVERS and GREENHOW, 1959). The absorption is caused by abnormal ionization in the lower ionosphere where the electromagnetic energy of the radio wave is dissipated by collisional processes. Night-time absorption takes place because of the presence of low-lying ionization generated by particles of high velocity.

The unique observations reported here of simultaneous emission and absorption in different sectors of the sky justify special attention since they rule out the possibility that the effects could be due to spurious causes and furthermore, indicate that both effects are caused by a common agent. All the absorption events occurred in the northerly sectors whilst emission events were recorded in all the sectors i.e. south, zenith and north. The northerly restriction of the absorption events indicates terminal features of charged particle streams in the lower ionosphere. The spatial distribution of the emission events indicates radiation generated by the transit through the upper ionosphere of the same particle streams which give rise to the absorption in the lower ionosphere near the auroral zones.

Examination of the available solar data indicates the occurrence of many minor flares during the period of unusual activity considered. Some of the events coincided with intervals of extended moderate solar disturbances, while others occurred about 1 hr after isolated flares. Although observationally interesting, these coincidences are statistically inadequate particularly as the other events appeared to be completely independent of any reported solar activity. The one solar feature which was prominent during the period 3–10 January 1959 was the appearance on 4 January of an unusually large sunspot with an area which reached over 3000 millionths of that of the solar disk and crossed central meridian on 10 January.

The lack of positive solar association with individual events makes it impossible to establish limits for the velocity of the corpuscular streams invoked. It seems probable that the unusual feature of the corpuscular streams producing the smooth effects of Fig. 1 would be their high velocity with energies in the cosmic ray region. This argument is supported by the facts that (a) the rare occurrence of these events, even at a time of relatively high solar activity, implies that energies in the normal corpuscular streams are inadequate, and (b) the theories for the generation of radiofrequency noise require high particle velocities as well as a substantial concentration of electrons. Such particle streams must have dimensions which are quite confined. Under the conditions featured in Fig. 1, the indicated separation between zones of emission and absorption does not exceed a few hundred kilometres assuming *F*-region emission and *E*-region absorption.

The IGY absorption programme has produced evidence bearing on emission of soft cosmic rays during solar flares. LEINBACH and REID (1959) have reported absorption records from Thule and two stations in Alaska following a solar flare giving strong suggestion of solar cosmic ray emission at low energies. BAILEY (1959) has examined the ionospheric and related effects of the great flare of 23 February

1956, and has concluded that the abnormal ionization in the lower ionosphere was associated with cosmic ray flux enhancements.

The *abrupt enhancements* in Fig. 2 have properties similar to solar noise bursts or even interfering signals. The possibility of interference is ruled out in most cases by (a) the frequency separation between receivers, (b) the isolated nature of the activity and (c) the associated magnetic disturbances. These events are thought to be closely related to the "auroral" noise reported by others. The signals were relatively strong and appeared in northern sectors only. Such events were essentially simultaneous in both northerly channels and persisted for less than 1 hr. The characteristic associated magnetic effect has the appearance of a highly damped pulse coinciding in time with the noise bursts. Most of the events occurred near local midnight but this could be a random coincidence.

Two mechanisms for producing these effects have been considered. The first assumes a process of noise generation in the outer atmosphere by the impact of high energy charged particles. If such is the case, however, one must recognize several obvious differences between the abrupt and the smooth enhancements. The magnetic effects indicate an abrupt change in conductivity, probably in the *E*-region, but no accompanying absorption was observed. The other mechanism could be unusual propagation of solar noise bursts around the earth from the sunlit hemisphere. The bursts in Fig. 2(a) for instance, occur at a time of "major radio noise bursts" in the sunlit hemisphere. The reported time of observation of 200 Mc/s radio noise immediately preceded the effect and the time of the 18 Mc/s noise bursts immediately followed. No reports of direct solar noise reception in the 80 Mc/s range are available, but it is evident that the event of Fig. 2(a) is bracketed in time and frequency by the other observations. Round-the-earth propagation of satellite signals on 40 Mc/s has already been reported by Wells (1958). Similar effects have been noted for 20 Mc/s satellite signals by a group at Stanford University (IGY progress reports) and for solar noise bursts at about 20 Mc/s by Smith *et al.* (1959). If the mechanism is propagation rather than generation, the simple image or focusing conception would not apply because of the pronounced northerly direction of arrival. In addition, the magnetic field effect implies a change in conductivity which cannot be a solar flare type of magnetic crochet since it occurs on the dark side of the earth.

The earlier observations at Jodrell Bank include examples of "auroral" noise at 80 Mc/s and those from Derwood, Maryland, show enhancements near local midnight associated with solar noise storms and magnetic disturbances. These are to be included in a further report. Meanwhile, the observing programme at Jodrell Bank is being continued in order to reach a better understanding of these unusual phenomena. In spite of the rare occurrence of such events, it is felt that the conclusive identification of radiofrequency noise generated in the earth's atmosphere will be of basic value and interest to studies of both solar–terrestrial relationships and radio-astronomy.

Acknowledgements—The writers wish to record the continued interest and support of Professor A. C. B. Lovell. They are indebted to Dr. C. D. Watkins for much of the reference material. The solar data was kindly supplied by the High Altitude

VOL.
17
1959/6

Observatory of the University of Colorado and by the Radiophysics Laboratory, C.S.I.R.O., Sydney, Australia.

References

Appleton E. V. and Hey J. S.	1946	*Phil. Mag.* **37,** 73.
Bailey D. K.	1959	*Proc. Inst. Radio Engrs., N.Y.*, **47,** 255.
Chapman S. and Little C. G.	1957	*J. Atmosph. Terr. Phys.* **10,** 20.
Chivers H. J. A. and Greenhow J. S.	1959	*J. Atmosph. Terr. Phys.* This issue, p. 1.
Covington A. E.	1947	*Terr. Magn. Atmos. Elect.* **52,** 339.
Ellis G. R.	1957	*J. Atmosph. Terr. Phys.* **10,** 302.
Hartz T. R.	1958	*Canad. J. Phys.* **36,** 677.
Hartz T. R., Reid G. C. and Vogen E. L.	1956	*Canad. J. Phys.* **34,** 728.
Leinbach H. and Reid G.	1959	*Phys. Rev. Letters* **2,** 61.
Smith A. G., Carr T. D. and Perkins W. H.,	1959	*Nature, Lond.* **183,** 597.
Wells H. W.	1958	*Proc. Inst. Radio Engrs., N.Y.* **46,** 610.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 20 to 33. Pergamon Press Ltd. Printed in Northern Ireland

The correlation of bursts of solar radio emission in the centimetre range with flares and sudden ionospheric disturbances

O. Hachenberg and A. Krüger
Heinrich-Hertz-Institut der Deutschen Akademie der
Wissenschaften zu Berlin, Berlin-Adlershof

(*Received* 30 *March* 1959)

Abstract—From statistical investigations of solar flares, of bursts of solar radiation in the centimetre range and of the accompanying effects of sudden ionospheric disturbances (SID) recorded during the first 6 months of the IGY, it follows that there exists a close correlation between bursts in the centimetre range and SID. Only those flares, which give rise to a centimetre burst, are also able to cause a SID. Both flares and the ionizing radiation of the ionospheric *D*-layer have their origin in the chromosphere. The close correlation confirms our view that the ionizing radiation as well as the centimetre radiation of the flares are generated by super-thermal electrons. The ionizing radiation is evidently a "Bremsstrahlung" of the X-ray region.

1. Introduction

There is no doubt that heavy outbursts of solar radio emission occur together with solar flares, as it was first stated by Appleton and Hey (1946). Also the short wave solar radiation in the ultra-violet or X-ray region, which seems to be responsible for the sudden ionospheric disturbances (SID), is produced in the same flare-phenomenon. But if individual flares are examined, some exhibit pronounced SID, others none at all; some flares give rise to pronounced bursts of radio emission, others do not. The connexion between these effects has not yet been satisfactorily explained.

This relation has been treated in several statistical investigations. Dodson *et al.* (1953), Davies (1954), Dodson *et al.* (1954), Loughhead *et al.* (1957), Budějický and Švestka (1958), Fokker (1958) have studied the problem.

With the exception of Dodson *et al.* (1954) on the one hand and Davies on the other, those authors based their studies mainly on measurements of burst radiation in the metre range ($\lambda > 0{\cdot}5$ m). Their results may be summarized as follows:

(a) It is certain that the flare areas of the sun often give rise to an additional radiation in the metre range.

(b) It is equally certain that not all flares produce an additional radiation. Some take place without being accompanied by a measurable burst at $\lambda > 0{\cdot}5$ m.

(c) X-radiation or u.v.-radiation may evidently be generated by flares producing an effect in the metre range as well as by flares not accompanied by any such burst.

It is difficult to obtain a more detailed description of these results by stating the percentage of the flares producing a burst at $\lambda > 0{\cdot}5$ m or that of the flares producing surely no such effect. Some authors give the percentages of flares accompanied by burst of single frequencies; their results are accumulated in Table 1.

The differences are due to the selection of the material used for statistics.

VOL.
17
1959/6

DODSON *et al.* (1953) associate all events at 200 Mc/s with a flare if the events were observed within an interval from 30 min before to 30 min after the flare. In our opinion these limits of association chosen are rather wide and are likely to contain a considerable amount of chance coincidences.

A similarly high percentage at 200 Mc/s was found by DAVIES. He further stated that by taking into account additional frequencies the percentage of coincidences can be further increased :

> But it seems hardly admissible to associate the type I peaks frequently occurring in the metre range with the flare phenomenon proper for the only reason of their coincidence.

Table 1. Percentage of flares accompanied by bursts

Source	Percentage of flares	Associated with bursts at
DODSON *et al.* (1953)	78	200 Mc/s
DAVIES (1954)	74	200 Mc/s
LOUGHHEAD *et al.* (1957)	23	Type III
FOKKER (1958)	14	200 Mc/s
BUDĚJICKÝ and ŠVESTKA (1958)	15·2	600 Mc/s

LOUGHHEAD *et al.* studied the correlation of type III bursts recorded by the radiospectrograph in the range 240–40 Mc/s with the flares observed in the same time interval. The restriction to type III bursts excludes the uncertain type I peaks and thus certainly avoids a great number of chance coincidences.

Also FOKKER adopted a rigid principle of selection and excluded type I peaks where as BUDĚJICKÝ and ŠVESTKA based their investigation only on 600 Mc/s measurements.

From the above it may reasonably be concluded that at least 14–20 per cent of the flares produce an effect in the metre range, while on the other hand it is certain that 25 per cent of the flares produce no effect at all. If we compare the coincidence of SID, we state that SID are associated with flares of the first group as well as with flares which do not produce any effect at $\lambda > 0{\cdot}5$ m.

It is generally supposed that the 200 Mc/s radiation originates in the corona, at a level about 50,000 km above the photosphere, whereas the optical flare is certainly a phenomenon of the chromosphere. Statistical analyses on the connexion between flares, bursts of the metre range and SID have not yet been very successful; this is mainly due to the fact that the effects of flare phenomena at entirely different levels of the solar atmosphere have been compared with one another.

If, on the other hand, we compare the centimetre radiation with the optical phenomena we may suppose, at least in the case of 3 cm, that this radiation originates in about the same gas volume as the optical emission. Besides, studies on the spectrum have shown (HACHENBERG, 1958) that the centimetre emission, contrary to that in the metre range, is a continuum. Statistics mainly based on 3·2 cm records are therefore representative of a much wider frequency range.

Post-burst increase and the subsequent monotonic decay of the radiation

indicate that at least part of the centimetre radiation is of quasi-thermic origin. For this reason we have to expect high electron velocities at the place of origin of the radiation.

The entirely different origin of the centimetre radiation as against the radiation in the metre range and the evident difference of the mechanisms of emission have led us to reconsider the relations between flares, centimetre radiation and SID.

2. The Data

The limits of detectability of the weak effects in question, the special methods of observation, and accidental errors of measurements restrict the material to a considerable degree and exert an extraordinary influence on the result of statistics. For this reason a careful discussion of the sources of error seems to be indispensable.

The present paper makes use of measurements made during the first 6 months of the International Geophysical Year (1 July to 31 December, 1957). The bursts of the centimetre region observed during that time were mainly taken from records of solar radiation at 9400 Mc/s made at Berlin-Adlershof. These bursts have been published in *Beobachtungsergebnisse des Heinrich-Hertz-Instituts;* in many cases, however, it proved useful to refer to the original records. In some cases records at $\lambda = 10$ cm obtained at Adlershof or Nera were used in addition. A total of 395 bursts was recorded in the period of observation; of these, 287 events were taken from 3·2 cm records, and, in addition, 108 were taken from 10 cm records.

In general the records allow a clear detection of enhanced radiation. The natural fluctuation of records averaged about ± 2 per cent of the daily radiation of the sun. The weakest bursts listed are those with intensities $J = 0{\cdot}02 J_{\odot}$, those with $J > 0{\cdot}04\, J_{\odot}$ could hardly escape observation. The lower limit of the recorded bursts is thus defined relatively clearly.

The conditions of the optical observation of the flares are not quite so favourable. For the daily intervals of observation in which there were clear records by the centimetre receivers, the reported flares were taken from the daily maps of the sun published by the Fraunhofer Institute at Freiburg. The observations were impaired by the weather, the image quality and, to a certain degree, by the observers themselves. All this resulted in some gaps in the material, which are somewhat smoothed out by data obtained from several other stations, but still make themselves felt. It is therefore possible to decide whether a reported flare was accompanied by a 3·2 cm burst; but on the other hand it cannot be stated with certainty whether an observed burst had its origin in a solar flare or not.

Concerning sudden ionospheric absorption (SID), identification is impaired by other influences. Though records of the field strength of distant short wave transmitters in the range between 2 and 6 Mc/s are very sensitive to small changes of the ionospheric absorption, they are extraordinarily influenced by fadings. Besides, the absorption depends on the angle of incidence of the solar radiation. The limit of identification of weak effects therefore depends on the time of day and the season. If instead of a short-wave transmitter the radiation of cosmic radio sources is recorded at 18–20 Mc/s, the records are free from fadings but also much less sensitive for the detection of weak SID.

VOL. 17 1959/6

The detection of SID is supported by records of atmospheric noise in the range 15 to 50 kc/s, which show the well-known "sudden enhancements of atmospherics" (SEA). The records give a relatively sensitive indication of enhanced electron density in the lower ionosphere and the related excessive absorption.

The present paper uses data of SID from the following stations: Heinrich-Hertz-Institut (Aussenstelle Neustrelitz), Observatorium für Ionosphärenforschung Kühlungsborn, Geophysikalisches Observatorium Collm Leipzig, Institut für Ionosphären-Physik Lindau, Observing Station Nederhorst den Berg-Radio (Nera).

Table 2. Numbers of flares, bursts and SID in the interval of observation

Month (1957)	Flares	Bursts 3·2 cm	Bursts 10 cm	SID
July	229	35	19	36
Aug.	243	40	23	38
Sept.	266	88	19	58
Oct.	318	64	21	48
Nov.	187	29	8	23
Dec.	153	31	18	28
	1396	total	395	231

In order to make sure that the reported effects were real and to avoid errors caused by long-period fadings or other local changes of ionospheric absorption, only those effects were included for statistics which had been stated by at least two observatories or had been obtained by the use of entirely independent methods. The figures for the observed events are compiled in Table 2.

3. The Relations Between Flares and Centimetre Bursts

According to Table 2, 1396 flares were observed and 395 bursts recorded in the time interval concerned. Of this number 280 bursts coincided with flares. Since the data of observation were selected in such a way that it was sure that there were 3·2 cm or 10 cm records during the time of the observed flares, it may be concluded that only 20 per cent of the flares are accompanied by an effect of intensity $J \geqslant 0{\cdot}04\ J_{\odot}$ in the centimetre range. With respect to the bursts this means that certainly 71 per cent of all bursts have their origin in the flare area.

If we suppose that it is due to unfavourable conditions of observation that no flares were seen in the case of the remaining 115 bursts, we should have to conclude that at least 74 per cent of the flares produce *no* effect with $J \geqslant 0{\cdot}04\ J_{\odot}$.

As on the other hand the number of bursts rises considerably with decreasing maximum intensity (Fig. 1), it has to be supposed that the percentage of coinciding events will increase as well with increasing sensitivity of the centimetre wave radiometers. If, for instance, we extrapolate the number of bursts down to the intensity $J = 0{\cdot}01\ J_{\odot}$ the number of coincidences will rise to about 30 per cent. The percentage of flares coinciding with bursts of certain intensity intervals will permit important conclusions on the average generation of energy in the flare area.

The distribution of the 280 bursts on the importance classes of the flares is shown in Fig. 2, curve I. Practically all flares of importance 3 and 3+ produce an effect in the centimetre range. On the other hand, out of thirteen subflares of

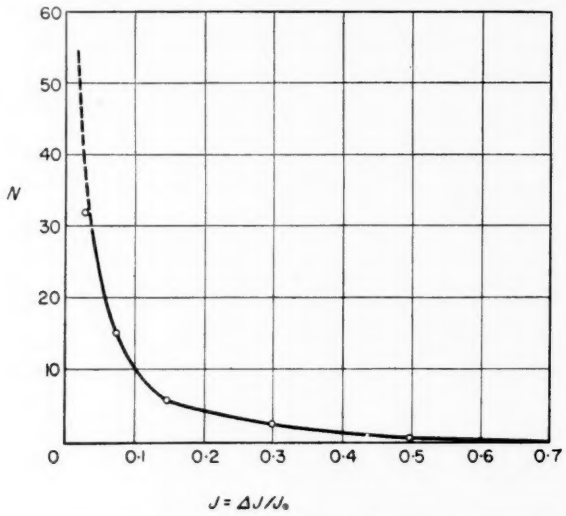


Fig. 1. Number of bursts against maximal intensity (N number in intervals of $\Delta I = 0{\cdot}01\ I_{\odot}$).

VOL.
17
1959/6

importance 1— only one is connected with a burst. Since as many as 115 bursts did not coincide with a flare and a fraction of 28 per cent of flares that escaped observation seems rather high considering the great number of observatories

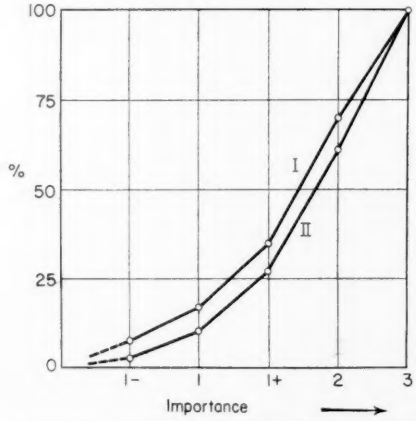


Fig. 2. Percentage of flares associated with bursts (I) and SID (II), in the total number of flares of different importances.

watching the sun, we are left with the possibility that some bursts start from weak 1— flares whose area is so little that they are near the limit of observation and thus likely to escape detection.

None the less we can summarize: The continuous emission in the centimetre

range is very closely connected with the processes in the flare region. But the relations between the optical flare area or their importance and the intensity of the centimetre radiation are only very loose ones.

4. The Relation Between Flares and Sudden Ionospheric Absorption (SID)

In the interval of observation, 231 cases of excessive absorption in the ionospheric *D*-layer (SID) were recorded for certain. If we compare their coincidence with the occurrence of flares, we find that 198 SID are associated with flares. With regard to all observed cases of SID this means that 85·7 per cent are certainly connected with flares. The high percentage confirms a result earlier obtained by Waldmeier (1948), who found that SID are directly connected with flare events.

If we suppose that SID always start from flares, the remaining 14·3 per cent must also be connected with flares, which, however, have escaped observation. On the other hand, we had to conclude from the relations between flares and bursts that 29 per cent of the flares must have escaped observation. The difference between these figures seems to support our view that part of the weak bursts is generated by very small subflares that cannot be observed because of their small area.

The percentage of flares of different importance connected with a SID is shown in Fig. 2, curve II. Apparently the flares of importance 3 and 3+ always produce a SID, while one out of ten flares of importance 1, and one out of thirty-six subflares of importance 1— produce an effect. The curve denoting the part of the flares in the different classes of importance connected with SID resembles that of the bursts but declines more markedly towards lower importances.

5. Relations Between SID and Centimetre Bursts

As already stated by Davies (1954), the sudden ionospheric disturbances are more likely to be accompanied by a burst in the centimetre region than by one at a lower frequency. This relation is shown clearly by the much more comprehensive material on which we could base this paper.

Of 231 SID, 210 were accompanied by bursts in the centimetre region. This means, that 91 per cent of all observed SID are associated with a centimetre burst. With regard to the 395 bursts the associations amount to 53 per cent.

The close association between SID and bursts of the centimetre region becomes evident, if we refer both effects to the solar flares. Out of 1396 flares, 280 are accompanied by bursts, out of which number 64 per cent also coincide with SID. On the other hand, there are 1116 flares which are surely not associated with bursts. At most twenty-one of these—or less than 2 per cent—coincide with SID. If we take into consideration the above mentioned difficulties to ascertain weak SID, we may conclude: only those flares which give rise to a burst in the cm region can also produce a sudden absorption by the ionosphere.

Now there arises the problem: are there any reasons why there were no SID in the other 185 cases? Let us begin by checking the question whether there exist certain types of bursts which are preferably associated or not with sudden

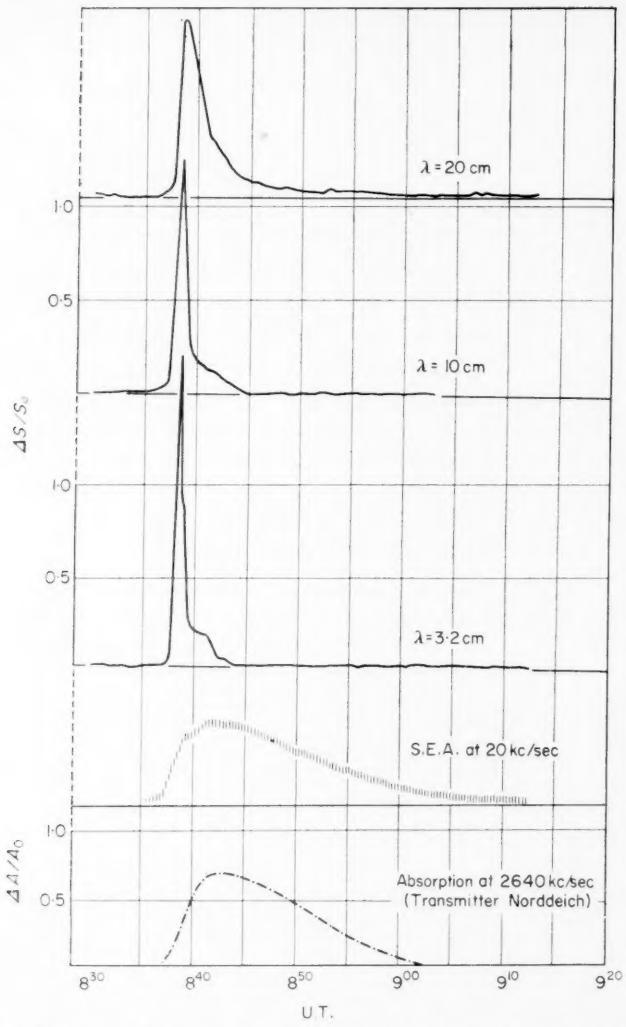


Fig. 3. Radio burst of 6 November 1957 in comparison with SEA and excessive absorption in the ionospheric D-layer.

absorption effects (SID). Starting from the classification by COVINGTON for burst records at 3000 Mc/s, we may distinguish four different types:

(a) Single peaks, the intensity of which rises to a maximum within a few seconds and then drops to the pre-burst level, however, a little more slowly.

(b) The "gradual rise and fall" of intensity taking place in time intervals of about 10 min or more.

(c) "Single peaks with post-burst increase"; after arising to a single peak similar to (a) the intensity drops to a level above pre-burst intensity. From there it decreases slowly to zero level in about 10 min or more.

(d) The complex burst composed of several maxima.

Reviewing the bursts recorded and comparing the corresponding records of field strength of distant short-wave transmitters, we find several specimen for every type of burst which are accompanied by a SID. Figs. 3–6 contain examples of different bursts, the curves of the accompanying ionospheric absorption and

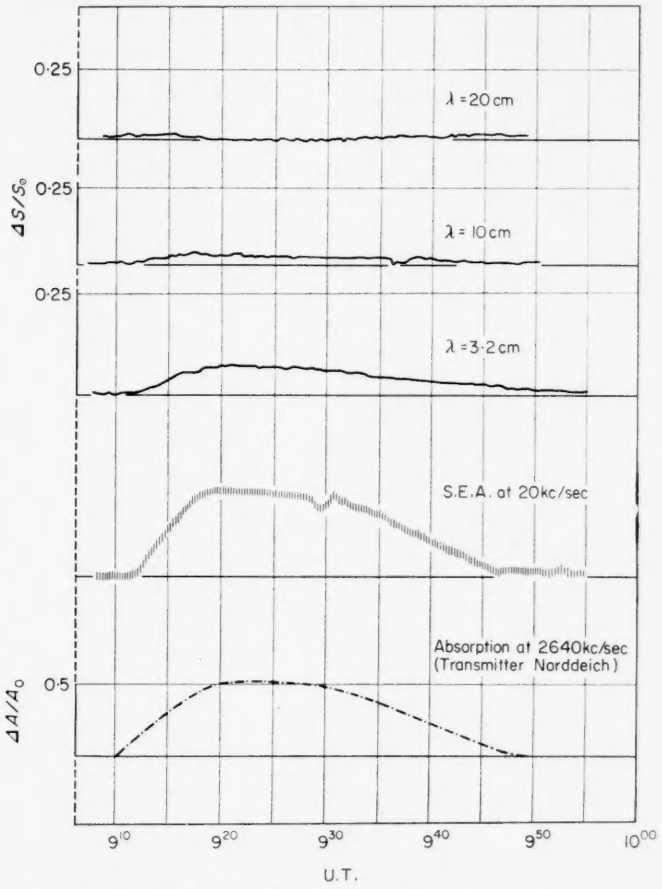


Fig. 4. Radio burst of 2 November 1957 in comparison with SEA and excessive absorption in the ionospheric *D*-layer.

those of the sudden enhancements of atmospheric noise. With one exception, the figures contain additional records representing the development of the bursts in the centimetre and decimetre ranges. Fig. 3 shows a burst of type (a) consisting of a single sharp peak with a half-width of 1·3 min. This burst is of high intensity also in the decimetre region; the maximum of SID is delayed by 4 min. Fig. 4 contains the record of a typical "gradual rise and fall"; this burst has a measurable intensity only at 3·2 cm, at 10 cm it is rather weak, and at 20 cm it produces no measurable effect at all. The corresponding curve of the ionospheric absorption runs nearly parallel to that of the 3·2 cm radiation, and their maxima coincide. Fig. 5 shows a peak with post-burst increase; in this case the absorption maximum

shows an 8 min delay compared with the peak of the burst. Lastly, a complex burst is shown in Fig. 6. The different shapes of the absorption curves represented in the above figures can be explained in the following way: we suppose the curve

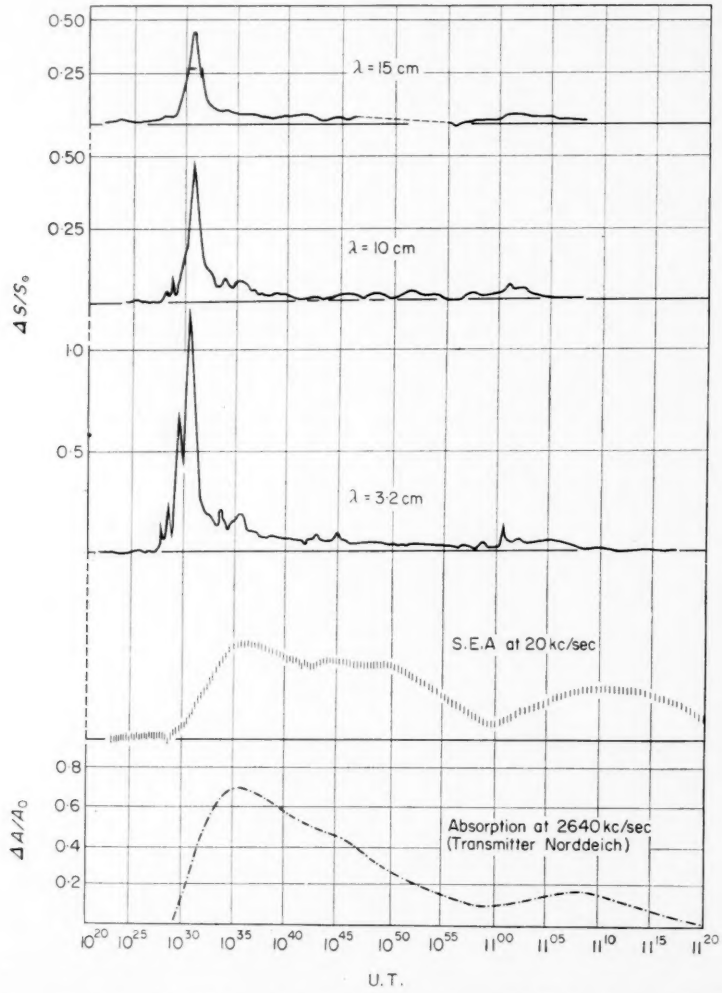


Fig. 5. Radio burst of 22 December 1957 in comparison with SEA and excessive absorption in the ionospheric *D*-layer.

of the ionizing radiation to be parallel to that of the 3 cm radiation, at the same time taking into account the delayed recombination of free electrons within the ionospheric *D* layer (HACHENBERG and VOLLAND, 1959). It is also worth mentioning that the curves of the SEA come very close to that of the absorption. Thus there is no indication of a type of burst principally not connected with a SID.

If, on the other hand, we study those cases in which a burst does not coincide with a SID, we find that we are primarily concerned with weak bursts; but among

VOL.
17
1959/6

these the type of single peak is prevalent. For the time being, this fact has to be regarded as a selection caused by the methods of observation, for among weak bursts short single peaks are much easier to detect than a very weak burst of the "gradual rise and fall" type.

Among the bursts that are not associated with a SID, one group deserves to be mentioned in particular: those bursts which are very weak at 3·2 cm, but extend strongly into the region of decimetre waves, and whose behaviour above

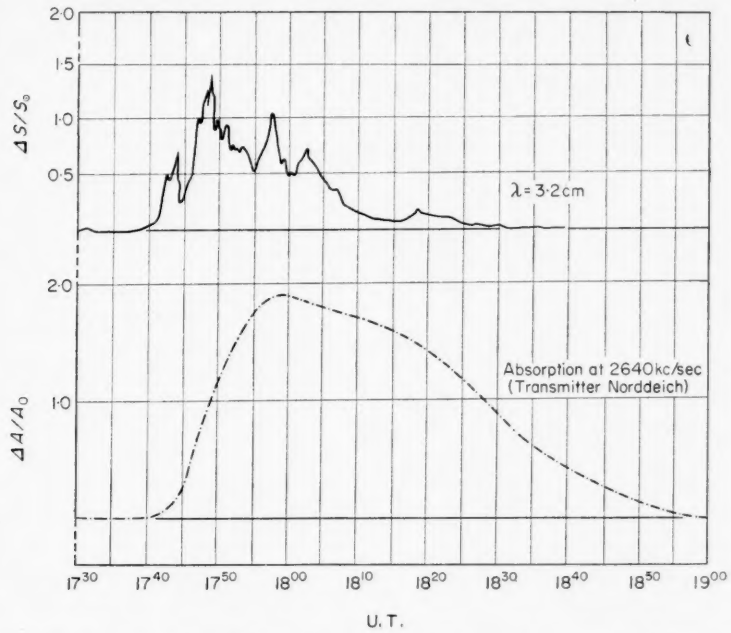


Fig. 6. Radio burst of 16 July 1957 in comparison with the excessive absorption in the ionospheric *D*-layer.

10 cm resembles already that of bursts in the metre region by their typical very sharp peaks. A burst of that type is represented in Fig. 7. Apparently this is one of the rare cases that a type III burst of the metre region can be traced into the region of centimetre waves; at least it is a burst preferably occurring in the corona.

As we said above, the material at hand indicates a certain dependence of the occurrence of SID on the intensity of the bursts. When we compared the time curves of "single peaks with post-burst increase" with the corresponding absorption curves, a relation between burst intensity and maximum absorption could be shown to be probable (Hachenberg and Vollard, 1959). A loose connexion between the total intensity and maximum absorption in the ionosphere is therefore not improbable.

In order to illustrate the influence of such a dependence on the results of statistics we measured the areas of the bursts, and hence derived the total of generated energy J_T. We preferred to use the total generated energy because of

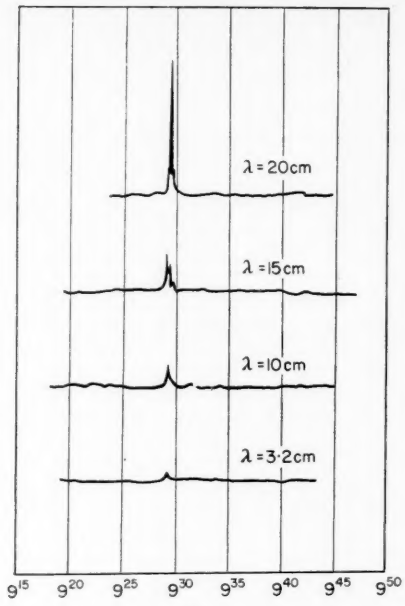


Fig. 7. Radio burst of 8 September 1957.

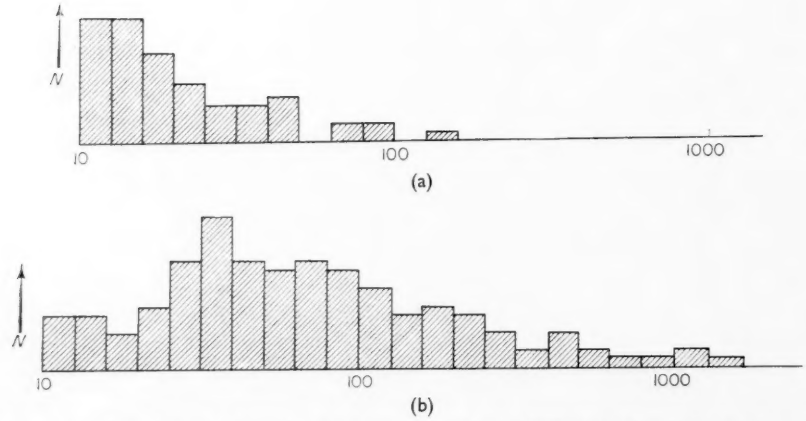


Fig. 8. (a) Number of 3·2 cm bursts accompanied by a SID plotted against total energy; (b) number of 3·2 cm bursts not accompanied by any SID.

the delay with which the electron density in the ionosphere follows changes in the ionizing radiation. If the number of bursts associated with a SID is plotted against the energy of the bursts, we obtain the histogram of Fig. 8(a). Inversely, the second part of the diagram records all those bursts which did not coincide with any detectable SID. It can be seen that all major centimetre bursts with

$$J_T > 200 \cdot 10^{-20} \text{ W-sec per m}^2\text{-c/s}$$

are clearly accompanied by a SID while the number of bursts associated with a SID drops at lower intensities. Contrarily, the number of bursts that did not show any noticeable SID rises considerably at low intensities. Whether the absence of SID in the case of weak bursts is caused by the limit of detection for weak effects, or whether SID do not occur together with weak bursts, cannot be decided for the time being.

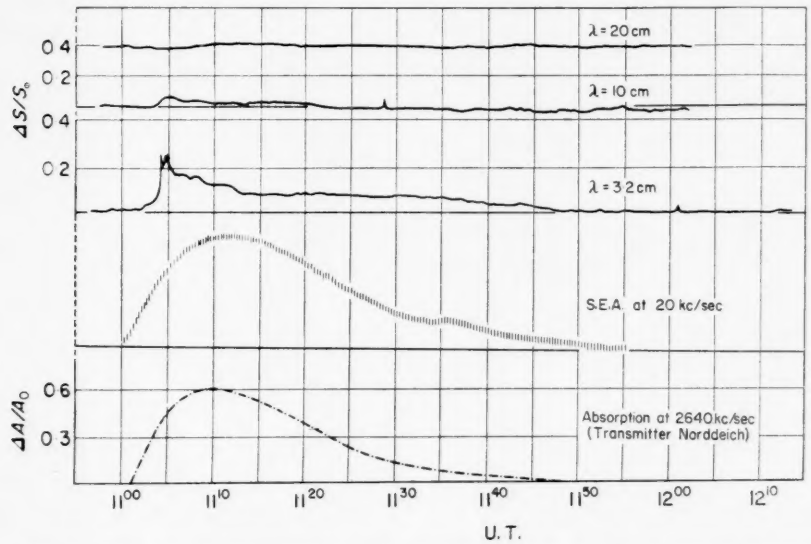


Fig. 9. Radio burst of 14 December 1957 in comparison with SEA and excessive absorption in the ionospheric *D*-layer.

6. Determination of the Height in the Solar Atmosphere Where the Ionizing Radiation Has Its Origin

From the close correlation between centimetre bursts and SID we can draw conclusions on the altitude in the atmosphere of the sun where the ionizing radiation has its origin. As already shown by Hachenberg (1958) there is a group of bursts exclusively taking place in the cm region and not extending into the dm and metre region. The ninety-five bursts of intensity $J_T \geqslant 100.10^{20}$ W-sec per m²-c/s include twenty-three of that kind. They are strong only at $\lambda = 3{\cdot}2$ cm; at 10 cm their intensity is very weak, and at $\lambda = 15$ cm or $\lambda = 20$ cm they are below the limit of detection. Bursts of that type are shown in Fig. 4 and Fig. 9. All of these twenty-three bursts are connected with a SID. If one supposes that these bursts do not appear above $\lambda = 20$ cm only because the emitting plasm is lying below the level at which $n = 0$ for $\lambda = 20$ cm, it follows that the region of emission must be less than 6000 km above the photosphere.

Contrarily, the bursts not associated with a SID include thirty-seven which take place predominantly in the decimetre and metre regions, like the example given in Fig. 7.

Within the interval of observation Potsdam and Nera finally reported 419 bursts at 200 Mc/s which did not coincide with phenomena in the centimetre

region, i.e., they occurred mainly in the corona. None of these bursts was accompanied by a SID.

Therefore the ionizing radiation generating the SID has its origin mainly in the region 2000 to 6000 km above the photosphere.

Conclusions

The statistics convinced us that the short-wave radiation which is responsible for the excessive absorption in the ionosphere occurs only when the emission of the continuum of the centimetre range is observed at the same time.

The ionizing radiation as well as centimetre radiation has its origin in the medium level of the chromosphere—in about the region 2000 km $< H <$ 6000 km—for the radiation is also associated with those bursts occurring only at 3·2 cm.

If the continuous emission of the cm region is interpreted as a quasi-thermal emission generated by free–free transitions, its occurrence indicates the existence of a highly accelerated electron gas—or super thermal electrons—in the plasm of the flare. Adopting the mean flare areas according to Waldmeier (1940), Ellison (1949) or Giovanelli (1958), one computes from $\lambda = 3{\cdot}2$ cm radiation temperatures of from 10^5 to 10^7 °K or mean electron energies from 1000 to 10^4 eV.

The super thermal electrons give rise to a "Bremsstrahlung" in the range of X-rays, the lower limit of which is between 10 and 1 Å.

The close correlation between centimetre radiation and the occurrence of SID supports our view that in the case of the ionizing radiation in the ionospheric *D*-layer we are concerned with just this "Bremsstrahlung", and not with a long wave monochromatic ultra-violet radiation, as one has for a long time been inclined to suppose. The recently published measurements with rockets by Chubb *et al.* (1958) may be regarded as a further indication; though there are some serious objections to these measurements as has been pointed out by Warwick and Zirin (1957).

Both X-radiation and centimetre radiation occur only in the case of sufficient acceleration of the electrons in the flare area. If the energy is $\ll 1000$ eV, it may still suffice to excite a flare in H_α, but it will no longer be enough to produce a SID or a centimetre burst. The frequent occurrence (79%) of flares without bursts and SID indicates that in the majority of flare phenomena the electron acceleration (or electron temperature) does not exceed the value critical for the X-ray component.

This idea may also serve to illustrate the results of earlier statistical analyses of bursts in the metre wave range compiled in the introduction. The bursts of the metre wave region which are not associated with phenomena in the centimetre wave region take place as monochromatic emissions of the plasm of the corona. They are not able to generate any X-radiation worth mentioning.

On the other hand those bursts of the metre wave range which are also strong in the centimetre wave region, give rise to an X-radiation which, however, has its origin in the chromosphere. This X-radiation causes a correlation of the bursts with SID. But there exists no direct relation between the phenomenon in the metre wave region and the SID, except that these bursts are excited by processes in the flare area.

VOL. 17 1959/6

References

Appleton E. V. and Hey J. S.	1946	*Phil. Mag. Series* 7, **37**, 73.
Budêjický J. and Švestka Z.	1958	*Bull. Astr. Insts. Csl.* **9**, 48.
Chubb T. A., Friedman H., Kreplin R. W. and Kupperian J. E., Jr.	1958	*Ann. Géophys.* **14**, 232.
Davies R. D.	1954	*Mon. Not. R. Astr. Soc.* **114**, 74.
Dodson H. W., Hedeman E. R. and Owren L.	1953	*Ap. J.* **118**, 196.
Dodson H. W., Hedeman E. R. and Covington A. E.	1954	*Ap. J.* **119**, 541.
Dodson H. W. and Hedeman E. R.	1958	*J. Geophys. Res.* **63**, 77.
Ellison M. A.	1949	*Mon. Not. R. Astr. Soc.* **109**, 3.
Fokker A. D.	1958	*Report and Document for I. A. U./U.R.S.I. Symposium on Radio-astronomy, Paris.*
Giovanelli R. G.	1958	*Aust. J. Phys.* **11**, 350.
Hachenberg O.	1958	*Z. Astrophys.* **46**, 67.
Hachenberg O. and Volland H.	1959	*Z. Astrophys.* **47**, 69.
Loughhead R. E., Roberts J. A. and McCabe M. K.	1957	*Aust. J. Phys.* **10**, 483.
Waldmeier M.	1940	*Z. Astrophys.* **20**, 46.
Waldmeier M.	1948	*Astr. Mitt. Zürich* No. 153.
Warwick J. W. and Zirin H.	1957	*Nature, Lond.* **180**, 500.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 34 to 45. Pergamon Press Ltd. Printed in Northern Ireland

The calculation of the m.u.f. factor for a non-parabolic ionospheric layer

M. D. VICKERS
DSIR Radio Research Station, Slough*

(*Received* 4 *April* 1959)

Abstract—A method is described for calculating the ray path of a radio wave through the ionosphere as represented by an $N(h)$ profile based on experimental data. A few such paths are calculated and from these m.u.f. factors are obtained. These factors are compared with those which would have been obtained had the existing methods of calculation been used. In most cases the differences are less than 4 per cent.

1. INTRODUCTION

MANY ionospheric observatories produce hourly values of the maximum usable frequency factor, (M 3000)*F2*, for reflections from the *F2*-layer over a range of 3000 km. One method used for the evaluation of (M 3000)*F2* depends on the assumption that the *F2*-layer has a vertical electron density distribution which is parabolic in shape. Another method is based on empirical relations derived from operational experience acquired over a long period of time.

Several research groups are now producing curves of electron density as a function of height in the ionosphere; these are referred to as $N(h)$ profiles. The calculation of these profiles includes no assumptions about the form of the electron-density distribution in the *F2*-layer or in any other layer, and it also takes into account the effect of the earth's magnetic field. These profiles often differ considerably from the familiar parabolic model of electron density distribution. In the field of radio communications, it is important to know what effect these differences have on calculations relating to oblique incidence propagation and, in particular, on the accuracy of published values of (M 3000)*F2*. This paper describes such an analysis for two typical profiles which differ considerably from a parabolic model.

2. PRESENT METHODS OF FINDING SKIP DISTANCE

2.1. *Theory*

A theoretical distribution of electron density as a function of height in an ionospheric layer was first worked out by CHAPMAN. Unfortunately the Chapman distribution is represented by a function, containing exponential terms, which is complicated to use analytically. For radio propagation near the maximum usable frequency (m.u.f.), however, only the part of the layer near the peak is important. To a good approximation, the Chapman distribution near the peak can be represented by a parabola.

APPLETON and BEYNON (1940, 1947) have shown that, by assuming the ionosphere to have a parabolic electron-density distribution, the equations for oblique-incidence radio ray paths in the ionosphere may be solved analytically. These equations give the propagation distance, D, for a single reflection from a layer in terms of the following quantities: h_m, y_m the height and semi-thickness of the

* Official communication.

equivalent parabolic layer; Δ, the angle of elevation of the ray above the horizontal; x, the ratio of signal frequency to critical frequency of the reflecting layer. Variations of y_m have little effect on D and may usually be neglected. Hence D can be expressed as a function of h_m, Δ and x.

To express D_s, the minimum value of D, as a function of x and h_m only, the equation for D is differentiated with respect to Δ. Equating the derivative to zero, solving for Δ, and inserting the resulting value in the equation for D gives D_s in terms of x and h_m. Alternatively, x may be expressed as a function of D and h_m; this alternative has been used by APPLETON and BEYNON to derive families of curves for selected values of D_s and h_m. In these relations, x is the maximum usable frequency factor for the distance D_s and to determine x, h_m must be known.

2.2. *Methods of deriving the height of a layer*

If certain approximations and simplifying assumptions are made, it is usually possible to define parabolic electron density distributions which represent independently the *E*-, *F1*- and *F2*-layers on any ionogram. Two of the methods which are generally used for finding the height of these equivalent parabolic layers are outlined below. When such an equivalent layer is assumed, any other layers which may be present are generally ignored in propagation problems.

2.2.1. APPLETON–BEYNON *method* (1940). For a parabolic layer the virtual height, h', is given, as a function of the vertical incidence sounding frequency, f, by:

$$h'(f) = y_m \frac{f}{f_c} \tanh^{-1} \frac{f}{f_c} + h_0$$

where h_0 = the height of the lower edge of the parabolic layer;
f_c = the critical frequency of the layer;
y_m = the semi-thickness of the layer.

If experimentally determined values of h' are plotted as function of $f/f_c \tanh^{-1} f/f_c$ the slope and intercept of a straight line fitted to the points will give y_m and h_0, respectively, for the equivalent parabolic distribution. The points will only lie accurately on a straight line if the electron-density distribution is parabolic but this is usually a good approximation in the range $0{\cdot}90 \leqslant f/f_c \leqslant 0{\cdot}99$ which is normally used. The height of the peak of the layer is obviously given by:

$$h_m = y_m + h_0$$

2.2.2. *Transmission curves* (Appleton–Beynon *theory*). The families of curves for D_s, x and h_m described in Section (2.1) can be used to construct a set of transmission curves in the form of a simple overlay for any type of ionogram. For any particular layer this overlay gives directly the maximum usable frequency for a distance of 3000 km. (M 3000) is obtained when this frequency is divided by the critical frequency of the layer considered. Knowing D and x, h_m is uniquely determined. Hence h_m can be obtained from (M 3000) by using the families of APPLETON–BEYNON curves in the reverse direction.

2.3. *Transmission curves (CRPL type)*

MARTYN's theorem states that if a radio wave of frequency f is reflected at vertical incidence at a particular height, then a wave of frequency $f \sec \phi$, incident on the ionosphere at angle ϕ, will be reflected at the same height. This theorem is only true in the special case of a flat earth. NEWBERN SMITH (1937) modified the theorem empirically to take the curvature of the earth into account.

Transmission curves are constructed from this empirical relationship and their use is described by the National Bureau of Standards (1948) and by WRIGHT *et al.* (1957). These transmission curves give distance of propagation as a function of Δ and x. Any particular constant distance transmission curve generally crosses the $h'(f)$ trace at two values of Δ corresponding to the high- and low-angle rays. In the particular case when the transmission curve is a tangent to the $h'(f)$ curve, the distance of propagation is the skip distance for the particular value of x considered. An alternative way of expressing this is to say that the given x is the m.u.f. factor for the given distance.

These CRPL curves have been found to be very similar to those derived from the parabolic model and will be considered here to be the same unless otherwise stated. It is important to note that the use of these CRPL transmission curves does not entail finding a value of h_m as an intermediate product in the calculation of distance of propagation.

3. ERRORS IN THE DETERMINATION OF SKIP DISTANCE

In the methods described above it is assumed or implied (CRPL method) that the *F2*-layer can be represented by a parabolic distribution of electron density with height. For this reason these methods will be referred to as "model methods". The height, h_m, of the apex of the parabola is chosen in such a way as to account for the observed change in the virtual height of the uppermost part of the *F2*-layer. Thus the effect of group retardation in any underlying layer that may be present is ignored and this often leads to values of h_m which exceed the true height, h_t, of the *F2*-layer by as much as 100 km.

If these two alternative values of layer height are inserted directly in the equation for D, the smaller value, h_t, will give the shorter distance of propagation for any given x and Δ. It might, therefore, be concluded that the larger m.u.f. factors, for a particular path, which would result from the use of h_t rather than h_m ought to be used in oblique incidence propagation problems. This conclusion is not, however, correct unless an allowance is made for refraction in the *E*- and *F1*-layers which, for a given signal frequency, increases the distance of propagation. MARTYN's theorem can be used to show that, for a flat earth, the values of skip distance, calculated using h_m without and h_t with refraction in the *E*- and *F1*-layers, are identical.

When the electron-density distribution is non-parabolic, the transmission curve technique ought to be a more reliable method of finding h_m than the straight line technique. This is because, when using the latter technique, a straight line has to be fitted to points which are not co-linear and large errors may be incurred as a result. Such errors are particularly likely to occur at equatorial stations

where the *F2*-layer has a more nearly linear than parabolic electron-density distribution.

4. Calculation of the Trajectory of a Radio Wave in a Layer of Arbitrary Distribution

4.1. *Theory*

The distribution of electron density with height is now calculated by various organizations using a method, originated by Budden (1954) and developed by Thomas *et al.* (1958), which makes no initial assumptions concerning the distribution. These distributions are usually referred to as $N(h)$ profiles. The calculation of a ray path through the ionosphere, as described by an $N(h)$ profile, is explained in the Appendix and the results are analysed in Section (5).

To simplify the calculation of the ray path, it was assumed (a) that ray theory is valid in the circumstances considered; (b) that the effect of the earth's magnetic field can be neglected at oblique incidence. Assumption (b) has been studied, under restricted conditions, by Haselgrove (1957) whose results suggest that, for distances and signal frequencies of practical application, the effect of neglecting the earth's field is small. It is, however, necessary to include the effect of the field in the calculation of the $N(h)$ profile.

A differential equation for the ray path is produced which reduces, in the case of a parabolic electron density distribution, to that derived by Appleton and Beynon (1940). This equation gives distance along the surface of the earth as a function of the height of penetration of the ray into the ionosphere. As the electron density represented by an $N(h)$ profile cannot be expressed in any convenient analytical form, the equation must be solved numerically.

The equation can be used to give either the actual trajectory through the ionosphere or the ground range, but only the latter has been evaluated.

4.2. *Results*

The distance of propagation, D, was evaluated as a function of the angle of elevation, Δ, for four constant values of x. Δ was taken from zero angle up to the maximum for which reflections from the *F2*-layer could occur. The resulting values of D were then compared directly with the values which would have been obtained if the model methods had been used.

Fig. 1(a) shows an $N(h)$ profile for summer noon at Slough when the *F1*-layer was prominent. Fig. 1(b) shows a Khartoum profile which was chosen because, even close to the peak of the *F2*-layer, it had a non-parabolic electron density distribution. For convenience, these will be referred to as profiles I and II, respectively. It should be emphasized that these profiles are extreme cases.

The height of the peak of the layers in profiles I and II, as determined using the various methods, is given in Table 1.

5. Propagation Characteristics Determined by Using Model and Profile Methods

5.1. *The skip distance*

Figs. 2 and 3 immediately give the skip distances for the given values of x. Table 2 gives skip distance as a function of x for profiles I and II.

Table 1. Height of peak of layer (km)

	Profile I		Profile II
Method	*F2*-layer	*F1*-layer	*F2*-layer
A	325	210	425
B	380	240	440
C	405	250	485

Method *A*: $N(h)$ profile method (Section 4.1)
B: Parabolic density distribution using a transmission curve (Section 2.2.2).
C: Parabolic density distribution using a straight line technique (Section 2.2.1).

Table 2. Skip distances (km) for profiles I and II

x	Profile I							Profile II			
	F2-layer				*F1*-layer			*F2*-layer			
Method	1·5	2·0	2·5	3·0	1·5	2·0	2·5	1·5	2·0	2·5	2·75
A	1160	1890	2840	4440	1180	1820	2850	1410	2280	3500	4700
B	1190	1880	2840	4800	1210	1860	3000	1400	2240	3450	4640
C	1260	1990	2980	5240	1220	1880	3050	1530	2480	3990	5880
D	1220	1920	2800	4000	1110	2000	3490	1380	2170	3150	3840

Method *A*: Deviation of the ray path through the ionosphere (Section 4.1).
B: Parabolic electron density distribution using a transmission curve method (Section 2.2.2).
C: As *B* but using a straight line technique (Section 2.2.1).
D: CRPL transmission curve method (Section 2.3).

Table 3. (M3000) *F2* for the various methods

	Profile I, *F2*-layer		Profile II, *F2*-layer	
Method	(M 3000) *F2*	Error (%)	(M 3000) *F2*	Error (%)
A	2·58	—	2·34	—
B	2·58	0	2·37	+1·5
C	2·50	−3	2·20	−6
D	2·61	+1	2·42	+3·5

VOL. 17 1959/6

For the *F2*-layer of profile I, there is no appreciable difference in skip distances whether determined using the $N(h)$ profile method (A) or the transmission curve methods (B and D). The distance determined using method C is about 8 per cent too high.

The skip distances for the *F1*-layer of profile I are not in such good agreement probably because the *F1*-layer critical frequency was not well defined on the original ionogram and had to be estimated. However, the differences are less

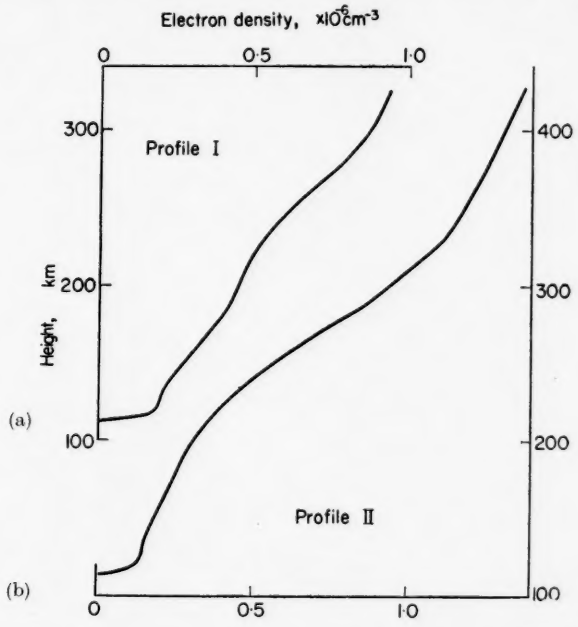


Fig. 1. Electron density profiles: (a) profile I; (b) profile II.

than 8 per cent which is comparable with the experimental errors involved in the various methods of calculation.

For profile II the values of skip distance for methods A and B are within a few per cent of each other. Method C, which uses a straight line technique to obtain the equivalent parabolic distribution, is inaccurate when applied to a non-parabolic layer and h_m may change by as much as 50 km when possible alternative straight lines are fitted to the chosen points (Section (2.2.1)).

The line actually chosen fitted the points best, but some empirical rule about the frequency range to be considered could probably be devised to ensure that the skip distances deduced in such circumstances were in better agreement with those obtained from the $N(h)$ profiles.

5.2. *The m.u.f. factor for* 3000 km

From Table 2, (M 3000) *F2* can be interpolated since it is the value of x for which $D_s = 3000$ km. (M 3000) *F2* is tabulated in Table 3 together with the percentage difference in factor between Method A and the other methods.

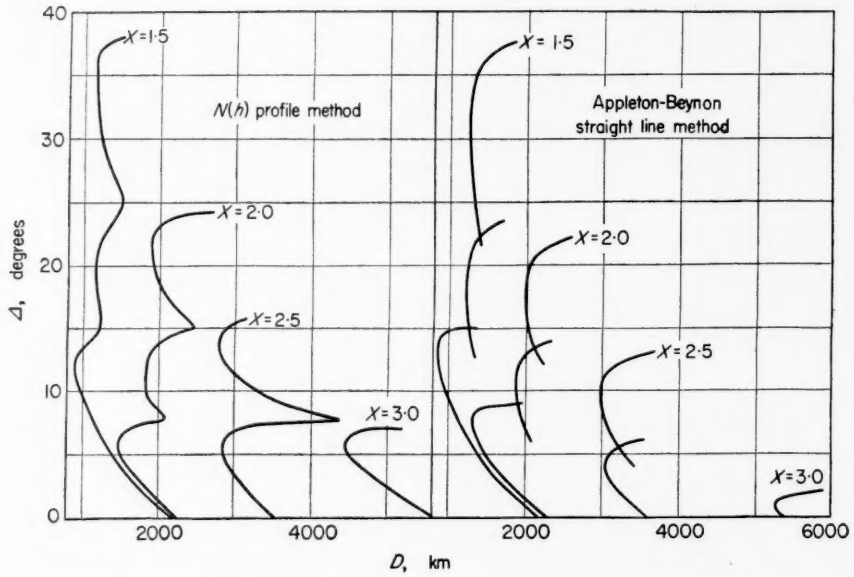


Fig. 2. The relation between angle of elevation (Δ) and distance of propagation (D).

VOL.
17
1959/6

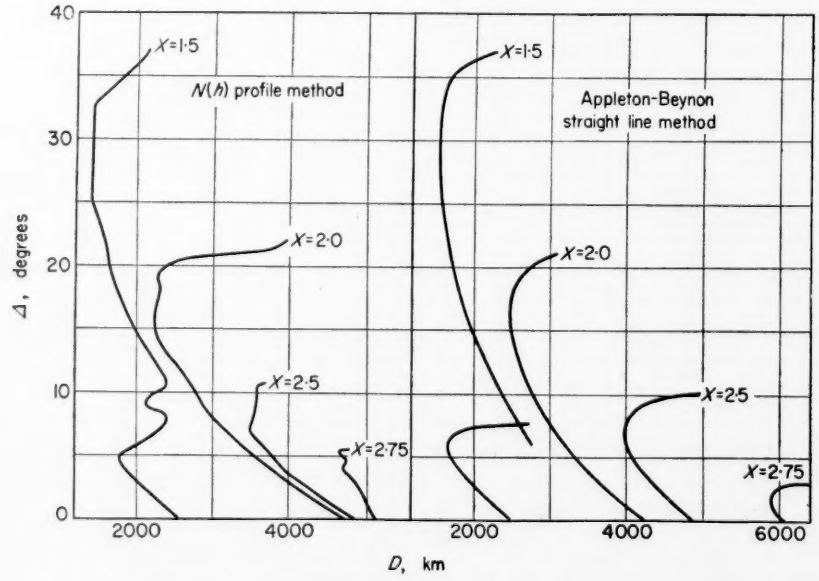


Fig. 3. The relation between angle of elevation (Δ) and distance of propagation (D).

5.3. *Angles of elevation*

5.3.1. *The angles for the skip distances.* At the skip distance for the *F2*-layer of profile I, the values of Δ_s derived using the model methods are lower than that obtained by the ray-tracing technique. For the parabolic layer methods, the difference varies from 5° to 3·5° as x increases from 1·5 to 3. For the CRPL curves, it is 2° for all values of x.

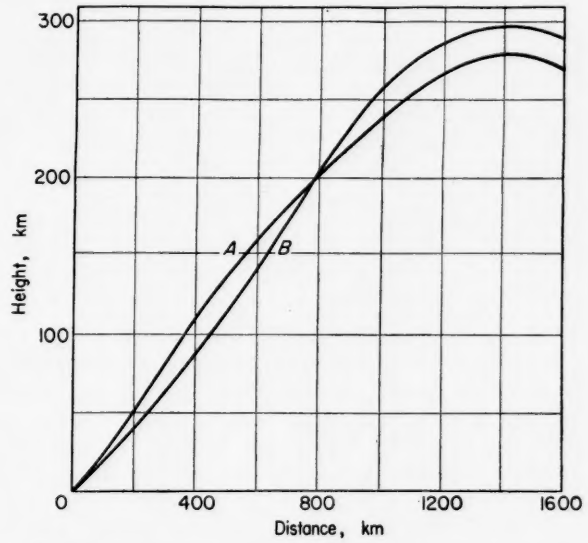


Fig. 4. Ray trajectories for profile I using methods *A* and *B*.

This difference in Δ_s for the same D_s is illustrated in Fig. 4 which represents the ray path through the ionosphere deduced, respectively, from *A*, profile I and *B*, the equivalent parabolic layer derived using a transmission curve.

Referring to Table 2 the value of skip distance for $x = 2{\cdot}5$ is the same for methods *A* and *B* and is 2840 km; the corresponding angles of elevation are 13·5° and 10·25°, respectively.

The apparent curvature at low levels in the trajectories arises because the heights above the curved surface of the earth are represented in cartesian co-ordinates.

The difference in ray path between the two methods of approach is clearly brought out in Fig. 4. The ray path for method *A* is continuously refracted between the height, at 100 km, when it enters the ionosphere and 280 km where it is reflected; whilst for method *B* refraction is confined between 240–300 km. The greater curvature of the ray path in the latter case occurs because the parabolic equivalent distribution for electron density in the ionosphere is a thin high layer.

5.3.2. *The penetration angle.* For a given value of x, the value of Δ corresponding to the penetration of the *F2*-layer is higher for the actual profiles than for any of the equivalent models. This is a similar effect to that described in the previous Section (5.3.1).

6. Discussion of Practical Implications

Assuming that the results described in Section (5) are generally true, the following conclusions regarding practical applications have been reached:

(a) The m.u.f. factors, and in particular (M 3000) *F2*, derived using present conventional methods are unlikely to be in error by more than 4 per cent.

(b) When the *E*- and *F1*-layers are present the angle of elevation of the ray, corresponding to the skip distance, may be 3 to 4° greater than that obtained using calculations based on a parabolic model.

This conclusion has an important bearing on the feasibility of single hop propagation to distances greater than 4000 km. Where previously it seemed that the only rays which could reach these distances were the "high-angle" rays, consideration of the trajectories through the actual ionospheric density distribution shows that low-angle propagation is possible at angles of elevation of about 3°.

7. Conclusions

Oblique incidence propagation characteristics are deduced, using a ray-path technique, for two typical distributions of electron density with height in the ionosphere ($N(h)$ profiles). These characteristics are compared with the results which would have been obtained had a parabolic electron-density distribution been assumed or had a CRPL transmission curve been used. The results of the comparison suggest that (a) skip distances derived using the two methods differ little from each other; (b) for a given skip distance, when the *E*- and *F1*-layers are present, the angle of elevation at the ground is higher for the $N(h)$ profile than for the model methods.

The method of deducing the ray paths from $N(h)$ profiles is simple, but rather tedious when carried out manually. Since only a few profiles were fully analysed, the conclusions above must be regarded as tentative. The two profiles examined here, however, were deliberately chosen to accentuate possible errors in deducing propagation characteristics which might arise using model methods.

Acknowledgements—The writer wishes to acknowledge the assistance given by Dr. C. M. Minnis for help in the preparation of this paper, and also that of Mr. F. Kift and Mr. E. D. R. Shearman whose helpful criticism made this paper possible.

The work described in the paper was carried out as part of the programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

Appendix

A1. *List of principal symbols*

D = propagation distance from transmitter to receiver;
f = transmitted frequency;
f_c = critical frequency of ionospheric layer;
$x = f/f_c$;
R = radius of the earth;
r_0 = radius of lower edge of the ionosphere;
r = radius to any height in the ionosphere;

VOL.
17
1959/6

y = height of ray path above the lower edge of the ionosphere;
y_m = height of the *F2*-layer max. above the lower edge of the ionosphere;
$Y = y/y_m$;
i = angle of incidence of a ray at a height r;
i_0 = angle of incidence of a ray at a height r_0;
Δ = angle of elevation at the ground;
μ = phase refractive index;
θ = angle subtended at the earth's centre by any two points on the ray path;
$N(Y)$ = electron density at any point in the ionosphere;
N_m = the peak value of the electron density;
f_N = plasma frequency for electron density N.

A2. *Calculation of the ray path*

The slope of the ray path in the ionosphere at any point is given by:

$$\tan i = \frac{r\,d\theta}{dr} \tag{1}$$

For a curved earth Snell's law is:

$$\mu r \sin i = \text{const} = r_0 \sin i_0 \tag{2}$$

where $\mu = 1$ at $r = r_0$

Hence
$$\tan i = \frac{r_0 \sin i_0}{(\mu^2 r^2 - r_0^2 \sin^2 i_0)^{1/2}}$$

and from (1)

$$\frac{dr}{d\theta} = \frac{r(\mu^2 r^2 - r_0^2 \sin^2 i_0)^{1/2}}{r_0 \sin i_0} \tag{3}$$

writing $r = r_0 + y = r_0\left(1 + \frac{y_m Y}{r_0}\right)$, and neglecting second order terms in y_m/r_0, which is much less than unity, equation (3) becomes:

$$\frac{{y_m}^2}{{r_0}^2}\left(\frac{dY}{d\theta}\right)^2 + 1 - \frac{\mu^2}{\sin^2 i_0} = \frac{2y_m Y}{r_0} + \frac{4y_m Y}{r_0}\left(\frac{\mu^2}{\sin^2 i_0} - 1\right) \tag{4}$$

but from (2)

$$\frac{\mu^2}{\sin^2 i_0} = \frac{{r_0}^2}{r^2 \sin^2 i}$$

For oblique-incidence propagation over distances greater than 1000 km, when the whole ray path is considered, the second term on the right-hand side of equation (4) is less than one fifth of the first term and may be neglected. This approximation is equivalent to the one used by Forsterling and Lassen (1931).

Equation (4) then becomes:

$$\sin {i_0}^2 \frac{y_m}{{r_0}^2}\left(\frac{dY}{d\theta}\right)^2 + \left(1 - \frac{2y_m Y}{r_0}\right)\sin^2 i_0 = \mu^2 \tag{5}$$

This is the differential equation of the ray path. If the earth's magnetic field is neglected,

$$\mu^2 = 1 - \frac{f_N^2}{f^2} = 1 - \frac{f_c^2 f_N^2}{f^2 f_c^2} = 1 - \frac{1}{x^2}\frac{N(Y)}{N_m}$$

and (5) becomes

$$\sin^2 i_0 \frac{y_m^2}{r_0^2}\left(\frac{dY}{d\theta}\right)^2 + \left(1 - \frac{2y_m Y}{r_0}\right)\sin^2 i_0 = 1 - \frac{1}{x^2}\frac{N(Y)}{N_m}$$

or

$$dY \Big/ \left\{x^2 \cos^2 i_0 + \frac{2y_m Y x^2 \sin^2 i_0}{r_0} - \frac{N(Y)}{N_m}\right\}^{1/2} = r_0 d\theta / x y_m \sin i_0$$

To calculate the horizontal distance to the reflexion point, the left-hand side of the equation is integrated from $Y = 0$ to $Y = Y_{\max}$. $Y_{\max}$ is the height of reflexion corresponding to the value of Y for which the denominator becomes infinite. Hence we have that:

$$\int_0^{Y_{\max}} dY / \{x^2 \cos {i_0}^2 + 2y_m Y x^2 \sin^2 i_0 / r_0 - N(Y)/N_m\}^{1/2}$$

$$= \int_0^{\theta} \frac{r_0 d\theta}{x y_m \sin i_0} = \frac{r_0 \theta}{x y_m \sin i_0} = \frac{r_0 D_1}{R^2} \cdot \frac{1}{x y_m \sin i_0} \quad (6)$$

where D_1 = ground distance for the ray in the ionosphere. For a parabolic electron density distribution

$$\frac{N(Y)}{N_m} = Y(2 - Y)$$

and equation (6) can be integrated analytically to give

$$D_1 = \frac{2R}{r_0} x y_m \sin i_0 \tanh^{-1}\left[x \cos i_0 / (1 - x^2 y_m \sin^2 i_0 / r_0)\right] \quad (7)$$

which is the same as the expression derived by APPLETON and BEYNON (1940).

A3. Numerical integration

The integral on the left-hand side of equation (6) was integrated using a seven-point Gauss integration formula. This method was chosen because it does not involve evaluating the integrand for $Y = Y_{\max}$ where the integrand becomes infinite. $N(Y)/N_m$ and $x^2 \cos^2 i_0 + 2y_m Y x^2 \sin^2 i_0 / r_0$ were plotted against Y. The point at which the two curves intersect gives the value of $Y_{\max}$. The Gauss sampling points were found as ratios of $Y_{\max}$, the corresponding values of the integrand were evaluated and the integral was calculated.

For convenience the height of the lower edge of the ionosphere was always taken as 100 km. The distance of propagation, D_2, below the ionosphere is calculated geometrically and is added to D_1 (equation 6) to give D, the total ground range of propagation.

VOL.
17
1959/6

References

Appleton E. V. and Beynon W. J. G.	1940	*Proc. Phys. Soc. Lond.* **52,** 518.
Appleton E. V. and Beynon W. J. G.	1947	*Proc. Phys. Soc. Lond.* **59,** 58.
Budden K. G.	1954	*Report of Cambridge Conference, on the Ionosphere*, p. 332. Physical Society. London.
Forsterling K. and Lassen H.	1931	*Z. tech. Phys.* **12,** 502.
Haselgrove J.	1957	*Proc. Phys. Soc.* B **70,** 653.
National Bureau of Standards	1948	*Ionospheric Radio Propagation.* Circular 462, 67–103.
Newbern Smith	1937	*J. Res. Nat. Bur. Stand.* **19,** 89.
Thomas J. O., Haselgrove J. and Robbins A. R.	1958	*J. Atmosph. Terr. Phys.* **13,** 46.
Wright J. W., Knecht R. W. and Davies K.	1957	*IGY Annals* **2,** Pt. I, 59.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 46 to 56. Pergamon Press Ltd. Printed in Northern Ireland

An analysis of a spectrogram of the aurora of 11 February 1958, in the wavelength range 3710-4420 Å*

L. WALLACE
Yerkes Observatory, University of Chicago

(*Received* 1 *April* 1959)

Abstract—A spectrum of the early stages of the great aurora of 11 February, 1958, covering the wavelength range 3710–4420 Å at a dispersion of 22 Å/mm shows several unusual features. The most striking of these are the lines of the $3d$–$4f$ transition array of *NII* which lie at excitation potentials 4·6 eV above those of the usually observed *NII* lines. The (0,0) and (0,1) bands of the N_2^+ first negative system indicate rotational temperatures of from 225°K for the lines with low K-numbers, to 575° for high K-numbers. The (2,12) and (1,12) Vegard–Kaplan bands of N_2 yield uncertain rotational temperatures with a lower limit of 800°K.

INTRODUCTION

A WELL-EXPOSED spectrum of the early stages of the bright type-*A* red aurora of 11 February, 1958,† in the wavelength range 3710–4420 Å was obtained with the 9 in. aperture $f/0{\cdot}8$ (Meinel) spectrograph at Yerkes Observatory. (This instrument has been described previously by CHAMBERLAIN (1955)). The exposure extended from 0200 UT, 11 February, about $\frac{1}{2}$ hr after the beginning of the aurora, until 0530 UT, 11 February, when the sky became heavily overcast.

The spectrum was photographed on an Eastman 103a-O plate through an optical step wedge but without the normal field lens and without a wavelength comparison. The third order of the grating was used to give a plate dispersion of about 22 Å/mm and an instrumental profile 1·5 Å wide at the half intensity points at λ3900.

IDENTIFICATIONS

The establishment of a wavelength scale for the plate was complicated by the lack of the normal comparison spectrum. The technique adopted was to assume certain generally accepted identifications (CHAMBERLAIN and OLIVER, 1953) and interpolate and extrapolate to obtain the wavelengths of the other features. As more identifications were made an increasingly better wavelength scale was established. The work of MOORE (1945) was used as the authority for atomic wavelengths and multiplet assignments, and that of PEARSE and GAYDON (1950) was used for the molecular band head wavelengths, except in the case of the Vegard–Kaplan bands where the measurements of HERMAN and HERMAN (1946) were used.

A list of identifications in order of wavelength made primarily on the basis of wavelengths and relative intensities is given as Table 1. This table is summarized

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)-3044.

† Solar, magnetic, X-ray, cosmic ray and all-sky camera observations associated with this aurora have been discussed by WINCKLER, *et al.* (1958).

Table 1. Identifications

λ observed	λ laboratory	Emitter
3709·7	3710·5	N_2 2P (2,4)
26·4	26·2	[O*II*], 1*
28·7	28·9	[O*II*], 1
49·5	49·5	O*II*, 3
54·4	55·4	N_2 2P (1,3)
67·2	67.	N_2 VK (2,12)
3804·1	3804·9	N_2 2P (0,2)
55·0	55.	N_2 VK (3,13)
57·1	57·9	N_2^+ 1N (2,2)
83·7	84·3	N_2^+ 1N (1,1)
†	3914·4	N_2^+ 1N (0,0)
3919·5	? 19·3	O*II*, 17
41·5	43·0	N_2 2P (2,5)
47·5	47·3	O*I*, 3
54·6	54·4	O*II*, 6
55·8	55·9	N*II*, 6
73·6	73·3	O*II*, 6
79·9	78.	N_2 VK (1,12)
82·9	82·7	O*II*, 6
95·2	95·0	N*II*, 12
97·3	98·4	N_2 2P (1,4)
4026·3	4026·1	N*II*, 40
35·4	35·1	N*II*, 39
41·5	41·3	N*II*, 39
43·6	43·5	N*II*, 39
46·6	46·6	Hg*I*, 1
58·5	59·4	N_2 2P (0,3)
70·1	69·9	O*II*, 10
72·7	72·2	O*II*, 10
76·1	75·9	O*II*, 10
79·3	? 78·9	O*II*, 10
92·9	? 92·9	O*II*, 10
4151·7	? 4151·5	N*I*, 6
71·8	71·6	N*II*, 43
76·3	76·2	N*II*, 42
85·4	85·5	O*II*, 36
89·7	89·8	O*II*, 36
4236·4	4236·5	N_2^+ 1N (1,2)
41·9	41·8	N*II*, 47, 48
†	78·1	N_2^+ 1N (1,0)
4317·7	4317·1	O*II*, 2
19·4	19·6	O*II*, 2
36·7	36·9	O*II*, 2
‡	40·5	Hγ
47·3	47·4	O*II*, 16
49·3	49·4	O*II*, 2
51·0	51·3	O*II*, 16
58·2	58·3	Hg*I*, 1
67·9	68·3	O*I*, 5
4414·6	4414·9	O*II*, 5
16·3	17·0	O*II*, 5

* Square brackets will be used to indicate forbidden atomic transitions. † Burned out. ‡ Diffuse.
The uncertain identifications are marked "?".
The mercury lines appear as contamination from street lights.

Table 2. Summary of identifications according to emitter

Atom	Multiplet number	Multiplet	Excitation potential (eV)
H		$H\gamma$	
NI	? 6*	$4s\ ^4P - 4p\ ^4S^\circ$	13·3
NII	6	$3s\ ^3P^\circ - 3p\ ^1D$	21·5
	12	$3s\ ^1P^\circ - 3p\ ^1D$	21·5
	39	$3d\ ^3F^\circ - 4f\ ^3G$	26·1
	40	$3d\ ^3F^\circ - 4f\ ^1G$	26·1
	42	$3d\ ^1D^\circ - 4f\ ^1F$	26·1
	43	$3d\ ^1D^\circ - 4f\ ^3F$	26·1
	47	$3d\ ^3D^\circ - 4f\ ^1F$	26·1
	48	$3d\ ^3D^\circ - 4f\ ^3F$	26·1
OI	3	$3\ ^5S^\circ - 4\ ^5P$	12·2
	5	$3\ ^3S^\circ - 4\ ^3P$	12·3
[OII]	1	$2p^3\ ^4S^\circ - 2p^3\ ^2D^\circ$	3·3
OII	2	$3s\ ^4P - 3p\ ^4P^\circ$	25·7
	3	$3s\ ^4P - 3p\ ^4S^\circ$	26·2
	5	$3s\ ^2P - 3p\ ^2D^\circ$	26·1
	6	$3s\ ^2P - 3p\ ^2P^\circ$	26·4
	10†	$3p\ ^4D^\circ - 3d\ ^4F$	28·6
	16	$3s'\ ^2D - 3p'\ ^2D^\circ$	28·4
	? 17‡	$3s'\ ^2D - 3p'\ ^2P^\circ$	28·7
	36	$3p'\ ^2F^\circ - 3d'\ ^2G$	31·2

Table 2 (*contd.*)

Molecule	Band system	Band
N_2	Second positive	(0,2), (1,3), (2,4), (0,3), (1,4), (2,5)
	Vegard–Kaplan§	(1,11)¶, (2,12), (3,13), (1,12), ? (2,13)*¶
N_2^+	First negative	(0,0), (1,1), (2,2), (0,1), (1,2)

* Located in a region of considerable blending.

† This identification is somewhat doubtful since it is not clear whether or not other multiplets of this transition array should be observed also.

‡ The brightest line of this multiplet, if present, is obscured by the (0,0) band of N_2^+.

§ The positions of the rotational lines were calculated from the rotational constants determined by CARROLL (1952) and LOFTUS (1956).

¶ Identified on the basis of rotational structure.

The uncertain identifications are marked "?".

and ordered according to emitter in Table 2. A densitometer tracing of the densest step of the spectrum is given as Fig. 1.

VOL. 17 1959/6

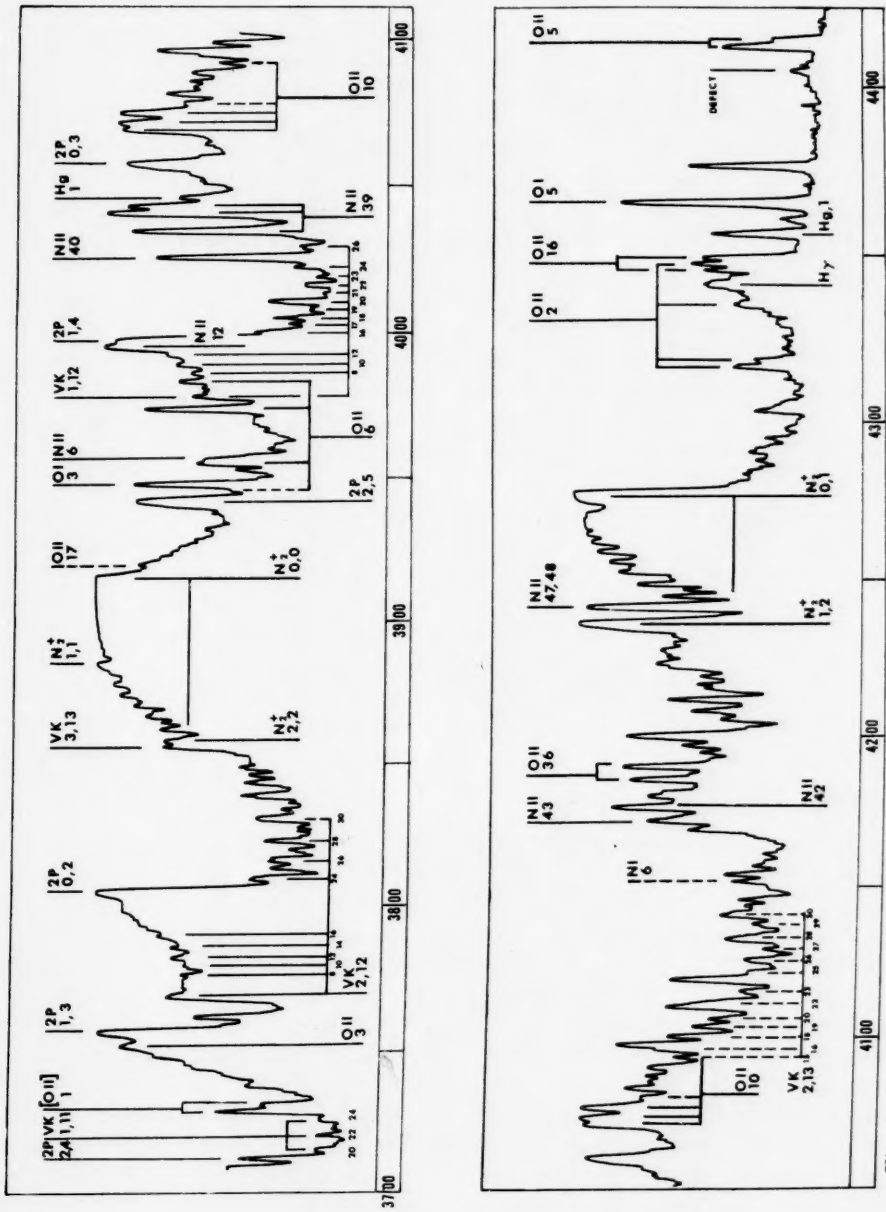


Fig. 1. Densitometer tracing of the spectrogram of the aurora of 11 February, 1958 with identifications and approximate wavelength scale. The identifications indicated by dashed leading lines are uncertain. Leading lines indicating the rotational lines of the Vegard–Kaplan bands are numbered according to their K-value.

DISCUSSION OF THE IDENTIFICATIONS

Most of the identifications—H, N*I*, O*I*, O*II*, N_2 second positive and N_2^+ first negative bands—only confirm those identifications which CHAMBERLAIN and OLIVER (1953) considered well established. The only outstanding feature is the $3d$–$4f$ transition array of N*II*. These lines have not been reported in the literature by the North American observers (BARBIER and WILLIAMS, 1950; PETRIE and SMALL, 1952, 1953; and VALLANCE JONES, 1955), but have been observed frequently by the Norwegians (VEGARD, 1950; VEGARD and KVIFTE, 1951; VEGARD and TØNSBERG, 1952; and possibly VEGARD and TØNSBERG, 1944). The appearance of this array with upper levels at 26·1 eV indicates a greater excitation than in the usual aurora observed in North America, in which the highest excited levels of N*II* are at 21·5 eV.

The resolution of the λ3727 feature into two features at wavelengths 3726·4 and 3728·7 tends to confirm the identification of this feature as the $\lambda\lambda$3726·2 and 3728·9 forbidden O*II* doublet. A single feature observed by OMHOLT (1957) at λ7325 and by DUFAY (1959) at λ7327 and a partially resolved feature at $\lambda\lambda$7320 and 7340 observed by DUFAY (1957) have previously been identified as multiplet 1 of [O*II*]. Confirmation of this identification is provided in the observed intensity ratio, $r = I(3729)/I(3726) = 0{\cdot}40 \pm 0{\cdot}04$ (m.e.). A theoretical expression developed for this ratio by SEATON and OSTERBROCK (1957) gives $r = 1{\cdot}5(1 + 2{\cdot}30x)/(1 + 9{\cdot}9x)$, where $x = 10^{-2}N_e/T_e^{\frac{1}{2}}$. The upper and lower limits of r from this expression are 1·5 and 0·35. Estimating $T_e \simeq 1000°$K and $N_e \simeq 10^6$/cm^3, a value of x is obtained ($\simeq$300) which could be thirty times lower and still give the theoretical lower limit for r of 0·35, which is in good agreement with the observed ratio.

Still further confirmation of the identification of the λ3727 feature as [O*II*] multiplet 1 is supplied by the identification of the [O*II*] multiplet 2 at $\lambda\lambda$7319 and 7330 on the Yerkes Observatory spectrum of the aurora of 5 December, 1958 obtained with the 9 in. aperture f/0·8 (Meinel) spectrograph. This removes the objection to the identification of the [O*II*] λ3727 doublet voiced by CHAMBERLAIN and OLIVER (1953). As suggested by CHAMBERLAIN and ROESLER (1955), the $\lambda\lambda$7319 and 7330 lines were observed superimposed on the lines of the (8,3) OH band, which occurs as night sky contamination. The appearance of these lines is illustrated in Fig. 2. The measured wavelengths of the forbidden lines were 7319 and 7331, and their relative intensities after correction for the OH contamination were $I(7319)/I(7330) = 1{\cdot}0 \pm 0{\cdot}2$. The theoretical high and low density limits of this ratio, 1·31 and 1·24, respectively, according to SEATON and OSTERBROCK (1957), tend to confirm the identification. A subsequent spectrum of the 11 January, 1959 aurora showed the (8,3) OH band contamination but not the forbidden lines.

Most of the unidentifiable features of the 11 February, 1958 spectrum fall in the region λ4080–λ4230. Some of the features are probably made up of blends of the rotational structure of the (2,13) and (3,14) Vegard–Kaplan N_2 bands and the (0,1), (1,2), and (2,3) first negative N_2^+ bands and possibly lines of the O*II* multiplets 19, 20 and 21 and the N*II* multiplet 38. Many more features than just these would be required to produce the observed spectrum.

The suggestion by SWINGS (1958) that the (0,9) Goldstein–Kaplan band of N_2

VOL. 17 1959/6

might be present in the auroral spectrum receives no support from the present observations, since the published wavelengths of the heads of this band do not correspond to any of the observed features.

The only unidentifiable strong lines falling outside the region λ4080–λ4230 are those observed at λ3760·0 and λ4378·8. The former, which does not appear to have been observed previously in the aurora, might be identifiable as multiplet 2 of *OIII*, but even though it would be expected to be the brightest line of *OIII* in the observed wavelength range, the identification of *OIII* on the basis of a single

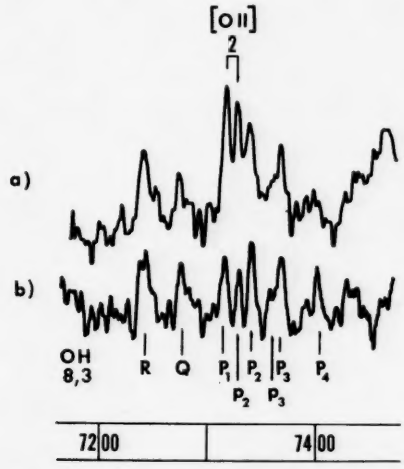


Fig. 2. Densitometer tracings of spectrograms of the region λ7200 to λ7420 of (a) the aurora of 5 December, 1958 and (b) the night sky. The night sky spectrum shows only the (8,3) OH band in this region while the auroral spectrum shows the forbidden multiplet 2 of [*OII*] overlapping rotational lines of the OH band.

line seems unwarranted. The λ4378·8 line appears to have been observed frequently by the Norwegian workers (Vegard and Tønsberg, 1941, 1944; Vegard, 1950; Vegard and Kvifte, 1951), but not by the North American observers. The most reasonable identification for this line would appear to be that of multiplet 17 of *NIII*. According to the published laboratory spectra (Fowler, 1919) this line might, with certain excitation conditions, be the brightest line of *NIII* in the observed wavelength region, and even though two of the lines of multiplet 1 of *NIII* might be identifiable as faint features on the present spectrogram, the arguments for identifying λ4378·8 as due to *NIII* are very weak.

Temperature Determinations

Characteristic curves were set up at various wavelengths of interest by using the relative transmission of the steps of the optical wedge through which the spectrum was photographed, and the corresponding densities of individual emission features. Intensity profiles of the (0,0) band of N_2^+ obtained from neighbouring steps showed about a 10 per cent scatter in the intensity measurements but no systematic trends.

The (0,0) and (0,1) first negative N_2^+ bands

Model profiles of the (0,0) and (0,1) N_2^+ first negative bands were calculated using the intensity relations and rotational constants given by HERZBERG (1950) and the rotational line wavelengths given by CHILDS (1932). The intensities of the model rotational lines were redistributed according to the observed instrumental profile. A comparison of the model profiles for rotational temperatures of 200°, 300°, 400°, and 600°K in the lower state, and the observed profiles is made in Fig. 3.

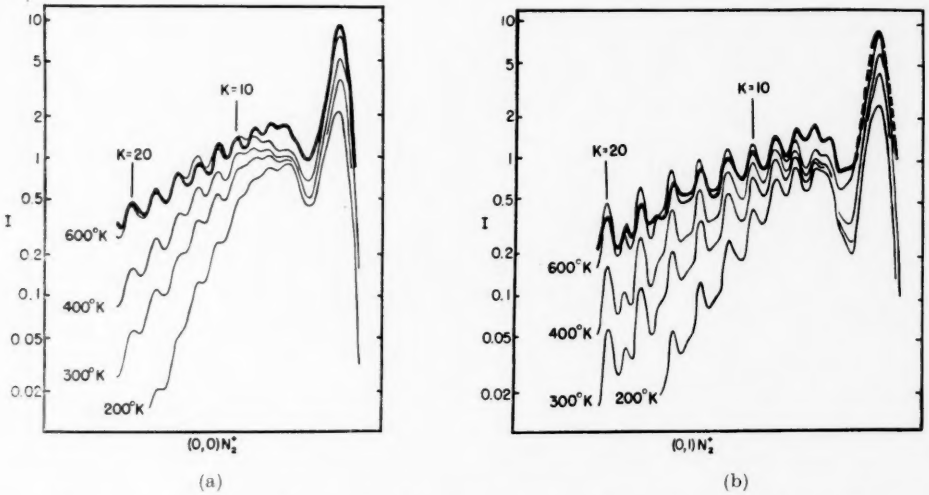


Fig. 3. A comparison of observed and model intensity profiles of the (0,0) and (0,1) N_2^+ first negative bands. The intensities are in arbitrary units. Light lines are used for the model profiles and dark lines for the observed. The dashed lines represent uncertain observations.

It is apparent that at least two temperatures are required to describe each of the observed profiles. From the shape of the intensity maxima in the tail, 250°K is obtained from the (0,0) band and 200°K from the (0,1) band. From the section of the tail for which $K \geqslant 10$, temperatures of 600°K and 550°K are obtained from the (0,0) and (0,1) bands, respectively. To within the accuracy of the photometry, the values for the two bands can be averaged to 225° and 575° with estimated uncertainties of 25 or 50°.

The 225° value is in reasonable accord with the values of from 200° to 350° usually observed (CHAMBERLAIN and MEINEL, 1954; VALLANCE JONES *et al.*, 1953; SHEPHERD and HUNTEN, 1955), but the apparent temperature increase with increasing K has only been reported by VALLANCE JONES and HARRISON (1955) who observed temperature increases of 100° going from small K to $K = 17$.

It seems unlikely that self-absorption could produce the observed effect, in the present observations at least, since the same temperature variation is observed in both the (0,0) and (0,1) bands. This mechanism would affect the (0,0) band much more than the (0,1) since the optical depth in the $v'' = 0$ level would be

much larger than in the $v'' = 1$ level. The suggestion that excitation by proton impact might produce the observed effect found no support in the laboratory bombardment experiments of ROESLER *et al.* (1958). The most appealing interpretation of the observed variation in the rotational temperature is that the bands reflect a true rotational temperature change with altitude.

The (2,12) *and* (1,12) *Vegard–Kaplan* N_2 *bands*

Model profiles were constructed for the (2,12) and (1,12) Vegard–Kaplan bands using the upper and lower state rotational constants determined by CARROLL (1952) and LOFTUS (1956), respectively. The line strengths used were those determined by SCHLAPP (1937). The models were calculated for upper state rotational temperatures of 800°K, 1600°K, and for the situation in which the Boltzmann factor $\exp[-(hc/kT)BK(K+1)]$ is unity for all values of K pertinent to the model. The central parts of the models were not calculated because the observed profiles are worthless in the region.

The observed profiles are of low quality because corrections were necessary in the regions near the heads for the overlapping second positive N_2 bands and in the tails for possible overlap with unknown features. In the regions near the heads, the observed profiles should be in error by no more than 10 or 20 per cent. In the tails, however, the background correction is completely uncertain. Consequently two profiles were determined for each band in this region—one for no background correction and one for maximum correction. The observed and model profiles are given in Fig. 4.

Allowing for a possible 20 per cent error in the observed head intensities, the (2,12) and (1,12) bands indicate rotational temperatures in the region from 850°K to 3400°K and 800°K to 4000°K, respectively. A reduction of the lower limits to 600°K would require the assumption of an additional factor of 2 error in the head to tail intensity ratios, while a reduction to 250°K would require a factor of 25 error. Since an additional factor of 2 error in the photometry of these bands seems out of the question, it appears necessary to accept a lower limit of 800°K.

The high rotational temperatures obtained from these bands tend to support the observation of high rotational temperatures from the N_2^+ bands.

DISCUSSION

The only outstanding features of the spectrum are: (a) the appearance of the $3d$–$4f$ transition array of *NII* originating from levels 4·6 eV higher than those usually observed in North America; (b) a range of rotational temperatures from the N_2^+ first negative bands of from 225 to 575°K; and (c) a lower limit on the rotational temperature of the Vegard–Kaplan bands of N_2 of 800°K.

When a considerable part of the emission comes from higher than usual altitudes—GARTLEIN and his associates (private communication) report upper limits of 800 km for this aurora—the appearance of higher than usual rotational temperatures is to be expected. Further, since there was no height discrimination in the present observations, one should not be surprised to observe molecular bands which could be interpreted as indicating a variation in rotational temperature.

A comparison of the present observations with those of CLARK and BELON (1959) and BELON and CLARK (1959) for the same aurora is difficult because their observations cover a large wavelength range at a dispersion of about 400 Å/mm, whereas the present observations cover a small wavelength range at about 22 Å/mm. Beyond the fact that both spectra are abnormal, there are few similarities. The abnormal intensities of the N_2^+ first negative bands in the sense that

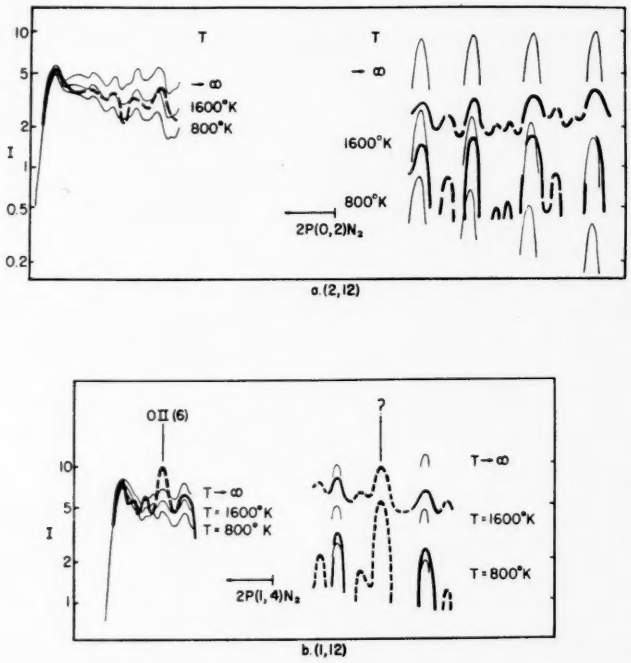


Fig. 4. A comparison of observed and model intensity profiles of (a) the (2,12) and (b) the (1,12) Vegard–Kaplan bands of N_2. The intensities are in arbitrary units. Light lines are used for the model profiles and dark lines for the observed. The dashed lines indicate uncertain observations. In the regions of the heads of these Vegard–Kaplan bands these result from uncertain background corrections and in the tails because the features are too faint for reliable photometry. In the case of the (1,12) band, the features marked "O*II* (6)" and "?" are clearly not parts of the band. The gaps in the middle of the profiles are the regions where the blending with the N_2 second positive bands is so severe that corrections for the presence of the N_2 second positive bands are impossible. Two profiles are given for the tail regions of the bands. The brighter is for no background correction and the fainter for correction such that the minima between the rotational lines are reduced to zero intensity. The correct tail profile should fall between these two extremes.

$I(0,2) : I(1,3) : I(2,4) = 100 : 130 : 60$ observed by CLARK and BELON can be compared with the essentially normal ratio which results from a conversion of the present observation of the $I(0,1) : I(1,2)$ ratio, by means of the transition probabilities of WALLACE and NICHOLLS (1955), to $I(0,2) : I(1,3) = 100 : 39$. The high rotational temperature of 2500°K found by CLARK and BELON for these bands is to be compared with a range of from 250° to 575° from the present observations. BELON and CLARK observe two multiplets of the $3d$–$4f$ transition

array which are outside of the present wavelength range, but do not observe other comparably strong lines of the array which are included in the present observations. Beyond this there is no similarity between the extraordinary atomic features reported by BELON and CLARK and those reported here in the wavelength region common to both investigations.

Acknowledgements—I would like to express my appreciation to Dr. J. W. CHAMBERLAIN for helpful discussions of many aspects of this work and to Mrs. V. PESCH and Mrs. F. VANDERVOORT for their photometric reductions of the spectrum and construction of the model intensity profiles. I would also like to thank Dr. C. W. GARTLEIN and his associates for supplying preliminary height determinations for this aurora.

REFERENCES

BARBIER D. and WILLIAMS D. R.	1950	*J. Geophys. Res.* **55,** 401.
BELON K. G. and CLARK A. E.	1959	*J. Atmosph. Terr. Phys.* **16,** 220.
CARROLL P. K.	1952	*Proc. Roy. Irish Acad.* A **54,** No. 22.
CHAMBERLAIN J. W.	1955	*Astrophys. J.* **121,** 277.
CHAMBERLAIN J. W. and MEINEL A. B.	1954	*The Earth as a Planet* (Edited by KUIPER G. P.). University of Chicago Press.
CHAMBERLAIN J. W. and OLIVER N. J.	1953	*J. Geophys. Res.* **58,** 457.
CHAMBERLAIN J. W. and ROESLER F. L.	1955	*Astrophys. J.* **121** , 541.
CHILDS, W. H. J.	1932	*Proc. Roy. Soc.* A **137,** 641.
CLARK A. E. and BELON K. C.	1959	*J. Atmosph. Terr. Phys.* **16,** 205.
DUFAY M.	1957	*C.R. Acad. Sci., Paris* **245,** 1648.
DUFAY M.	1959	*Ann. Geophys.* **15,** 134.
FOWLER A.	1919	*Mon. Not. R. Astr. Soc.* **80,** 692.
HERMAN R. and HERMAN L.	1946	*J. Phys. Radium* **7,** Series 8, 203.
HERZBERG G.	1950	*Spectra of Diatomic Molecules* (2nd Ed.) Van Nostrand, New York.
LOFTUS A.	1956	*Canad. J. Phys.* **34,** 780.
MOORE C. E.	1945	*A Multiplet Table of Astrophysical Interest.* Contribution Princeton University Observatory, No. 20.
OMHOLT A.	1957	*J. Atmosph. Terr. Phys.* **10,** 320.
PEARSE R. W. B. and GAYDON A. G.	1950	*The Identification of Molecular Spectra* (2nd Ed.). Chapman and Hall, London.
PETRIE W. and SMALL R.	1952	*Astrophys. J.* **116,** 433.
PETRIE W. and SMALL R.	1953	*Canad. J. Phys.* **31,** 911.
ROESLER F. L., FAN C. Y. and CHAMBERLAIN J. W.	1958	*J. Atmosph. Terr. Phys.* **12,** 200.
SEATON M. J. and OSTERBROCK D. E.	1957	*Astrophys. J.* **125,** 66.
SCHLAPP R.	1937	*Phys. Rev.* **51,** 342.
SHEPHERD G. G. and HUNTEN D. M.	1955	*J. Atmosph. Terr. Phys.* **6,** 328.
SWINGS P.	1958	*Technical Final Report.* Contract AF 61(514)-912, University of Liege, Cointe-Sclessin, (Belgium).
VALLANCE JONES A.	1955	Scientific Report No. AR-20. University of Saskatchewan, Saskatoon.
VALLANCE JONES A. and HARRISON A. W.	1955	*J. Atmosph. Terr. Phys.* **6,** 336.
VALLANCE JONES A., HUNTEN D. M. and SHEPHERD G. G.	1953	*Astrophys. J.* **118,** 350.

VEGARD L.	1950	*C. R. Acad. Sci., Paris* **230,** 1884.
VEGARD L. and KVIFTE G.	1951	*Geofys. Publ.* **18,** No. 3.
VEGARD L. and TØNSBERG E.	1941	*Geofys. Publ.* **13,** No. 5.
VEGARD L. and TØNSBERG E.	1944	*Geofys. Publ.* **16,** No. 2.
VEGARD L. and TØNSBERG E.	1952	*Geofys. Publ.* **18,** No. 8.
WALLACE L. V. and NICHOLLS R. W.	1955	*J. Atmosph. Terr. Phys.* **7,** 101.
WINCKLER J. R., PETERSON L., HOFFMAN R. and ARNOLDY R.	1958	Technical Report 2.12, University of Minnesota, Minneapolis.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17: pp. 57 to 64. Pergamon Press Ltd. Printed in Northern Ireland

A daily index of solar activity based on *E*-layer ionization (July 1957–December 1958)

C. M. Minnis and G. H. Bazzard*
DSIR Radio Research Station, Slough

(*Received* 10 *April* 1959)

Abstract—The electron density in a Chapman layer can be related to the intensity of the incident solar ionizing radiation. This relation has been adopted as the basis for computing a daily index of the radiation intensity using the critical frequency of the *E*-layer at Slough. Precautions have been taken to minimize errors due to irregularities in the behaviour of the layer and to the difficulty in identifying the *E*-layer cusp. The standard deviation of the residual errors in the index is estimated to be 2 per cent. The index has been tabulated for the period 1 July 1957 to 31 December 1958.

1. Introduction

During the next few years, considerable effort will be devoted to the study of geophysical data of many kinds which have been collected during the IGY. In certain fields, such as geomagnetism and the ionosphere, solar radiation has an important controlling influence and, when analysing the experimental data, it is often desirable to have some measure of the intensity of this radiation.

The active component of the radiation is absorbed in the upper atmosphere and cannot be directly observed at ground level. Fortunately there is a fairly reliable relation which connects the degree of ionization in the *E*-layer with the intensity of a particular component of the incident radiation; this is the component which gives rise to the *E*-layer and which probably lies in the X-ray region. Since the critical frequency, and hence the degree of ionization, in the *E*-layer can easily be measured, the use of the relation just mentioned suggests a possible basis for an index of solar activity intended to represent the changes in the intensity of the ionizing radiation.

A monthly mean index of this kind has already been tabulated by Minnis and Bazzard (1959) but, as a result of proposals made by the URSI–AGI Committee (1958), an attempt is made in the present paper to produce a daily index, based on *E*-layer measurements, for the period 1 July 1957–31 December 1958.

2. Theoretical Considerations

In a Chapman layer, the electron density, N, at a fixed height is given by the equation:

$$\frac{dN}{dt} = q_0 \cos^n \chi - \alpha' N^2 \qquad (1)$$

where

q_0 = rate of production of electrons for $\chi = 0$;
α' = effective recombination coefficient;
χ = zenith angle of the sun;
$n = 1$.

* Official communication.

If the diurnal change in the height of the peak of the layer is small, it may be assumed that N represents the electron density corresponding to the critical frequency of the layer. Equation (1) may then be written

$$f^4 = \left(\frac{q_0}{K\alpha'} \cos^n \chi\right)\left(1 - \frac{1/\alpha' \,.\, dN/dt}{(q_0/\alpha') \cos^n \chi}\right) \tag{2}$$

where

$K = 1{\cdot}54 \times 10^8$

f = critical frequency (Mc/s)

Except near sunrise and sunset when dN/dt is large, it follows from equation (2) that

$$f \propto \cos^{n/4} \chi \tag{3}$$

$$f^4 \sec^n \chi \propto \frac{q_0}{\alpha'} \tag{4}$$

In the absence of any disturbing influences, q_0 is proportional to the flux, S, of the solar ionizing radiation. Hence if α' remains constant,

$$f^4 \sec^n \chi \propto \frac{q_0}{\alpha'} \propto S \tag{5}$$

It is reasonable, therefore, to use the critical frequency data for a Chapman layer as the basis for computing an index number which is intended to be proportional to the intensity of the ionizing radiation.

For an ideal Chapman layer, $n = 1$ in equation (3) and hence $f \propto \cos^{\frac{1}{4}} \chi$ when dN/dt is small, but it is often found that the experimental value of the exponent of cos χ differs from $\frac{1}{4}$. Nicolet and Bossy (1949) have shown that vertical gradients in the scale height of the atmosphere could lead to such discrepancies. Beynon and Brown (1956) have suggested that the value of the exponent may also depend on geomagnetic influences. In view of the divergence between the actual and theoretical values of n, it is essential, in any application of equation (5), to use the appropriate experimental value of n which may not be unity.

3. *E*-Layer Indices of Solar Radiation

3.1. *Annual and monthly mean values*

Appleton and Naismith (1939) used the annual mean values of $Ch_E = (f_0E)^4 \sec \chi$, the "character figure", to show that this index varied in phase with the solar cycle. Monthly mean values of Ch_E for the years 1931–1946 are shown graphically in a report issued by the DSIR (1953). For the years 1938–1958, Minnis and Bazzard (1959) have reproduced tables of monthly mean values of an index, I_E, which is related to Ch_E. In all these cases, the values of Ch_E were calculated using the Slough noon critical frequencies and assuming $n = 1$; this value for n is justified because the seasonal variations of f_0E at noon at Slough are found to be proportional to $\cos^{\frac{1}{4}} \chi$ even though the diurnal variations are proportional to $\cos^{\frac{1}{4}} \chi$.

VOL. 17 1959/6

3.2. *Accuracy of mean values of* $\mathrm{Ch_E}$

The monthly mean values of Ch_E used by MINNIS and BAZZARD in deriving I_E were based on the daily routine tabulations of f_0E at noon. It is usual to read f_0E to the nearest 0·1 or 0·05 Mc/s depending on the quality of the ionogram. Since each hourly reading is made independently, it is possible that the tabulated values may not always refer to precisely the same stratum of the E-layer. The $E2$-cusp, when present, can usually be distinguished, but subsidiary stratifications in the lower part of the layer often make it difficult to decide on the best value for f_0E particularly when E_s-ionization is present below the peak of the layer. Hence, in addition to rounding-off errors of up to $\pm$0·05 Mc/s, identification errors of $\pm$0·10 Mc/s, or more, may also occur in the normal tabulations of f_0E. Errors of this magnitude would result in spurious fluctuations of $\pm$10–15 per cent in the daily values of Ch_E. When monthly mean values of Ch_E are computed, the fluctuations due to this cause fall to about 2 per cent and, assuming that there are no systematic errors in the readings, they are not important for many purposes.

3.3. *Daily values of* $\mathrm{Ch_E}$

In the preceding Section, it was shown that if the normally tabulated noon values of f_0E were used as the basic data, the resulting daily values of Ch_E would often contain day-to-day fluctuations of $\pm$10–15 per cent due to errors in the identification of the correct cusp and to rounding-off errors. In view of this, it appears that in any attempt to compute accurate daily values of Ch_E, two precautions must be taken: (a) f_0E must, wherever possible, be read more accurately than is normal in routine tabulations; (b) the mutual consistency of all the hourly values for each day must be examined to ensure the correct identification of the E-layer cusp.

4. THE DAILY INDEX J_E AT SLOUGH

4.1. *The experimental data*

The daily values of the index J_E in Tables 1 and 2 were derived after a re-examination of all the relevant ionograms, in the course of which the precautions mentioned above were carefully observed. The frequency marks on the Slough ionograms usually allow f_0E to be read to the nearest 0·01 or 0·02 Mc/s when the cusp is well defined. Each of the hourly ionograms from 0800 to 1600 LMT inclusive was examined and the values of f_0E were read to this accuracy wherever possible. For each day, a plot was made of $\log f_0E$ against $\log \cos \chi$ and these log–log plots formed the basic data for subsequent analysis.

On many days, most of the points lie, as expected, close to a line defined by

$$\log (f_0E) = \tfrac{1}{3} \log (\cos \chi) + \text{constant} \qquad (6)$$

which is the appropriate statistical relation for the diurnal variation in f_0E at Slough. It is common, however, to find that some points do not lie on this line. Some of the values of f_0E which are too high have been found to coincide with solar flares which have been reported. The values of f_0E near sunrise and sunset tend to lie, respectively, below and above those indicated by the line, but this departure is in accordance with the term containing dN/dt in equation (2).

Although some of the departures from linearity in the daily log–log plots can be accounted for, there are other occasions when no explanation can be found and when it must be assumed that some unidentified disturbance has led to a temporary breakdown in the normal balance between the production and disappearance of electrons.

4.2. *Derivation of* J_E

In view of the irregularities just described, it is obviously dangerous to assume that any single measured value of f_0E can be used to derive a value of Ch_E which accurately represents the level of solar activity on the day in question. Some method of taking into account all the measured values of f_0E for each day is essential.

If a straight line, with the appropriate slope, is drawn on one of the log–log graticules mentioned in Section (4.1), it may be assumed to represent the diurnal changes in f_0E for some constant value, S, of the intensity of the ionizing radiation. If this line intersects the ordinate corresponding to the noon value, χ_0, of the solar zenith angle at a frequency f_r, the quantity J_E defined by

$$J_E = f_r^4 \sec \chi_0 \tag{7}$$

may, with the reservations referred to in Section (2), be assumed to be proportional to S. When the observed values of f_0E for a given day lie on or very close to such a line, the value of f_r can easily be determined and inserted in equation (7) to give the characteristic value of J_E for the day.

It is more usual to find that the points do not lie accurately on a straight line and it then follows that each point by itself could be assumed to represent a different value of S or of J_E. To overcome this difficulty, the technique adopted in the determination of the data in Tables 1 and 2 was to draw, visually, the line which best fitted the observed values of f_0E and then to read the value of f_r defined by this line. The insertion of this value of f_r in equation (7) leads to a value of J_E which may be assumed to represent the hypothetical constant value of S defined by the line which, in turn, best represents the observed values of f_0E for the day.

4.3. *Drawing the best-fit line*

The values of f_0E nearest sunrise and sunset seem to be more prone to unaccountable fluctuations than those nearer noon. When drawing the lines, greater weight was given to the points nearer noon and also to the more reliable values of f_0E ($\pm 0{\cdot}02$ Mc/s). When the error in determining f_0E seemed likely to exceed 0·05 Mc/s, the point was not included in the log–log plot.

On days when, for any reason, the points available were limited in number and of doubtful accuracy, the line drawn through the points was assumed to have a slope corresponding to $f \propto \cos^{\frac{1}{4}} \chi$ which is the appropriate statistical relation for Slough.

Although in fact the forenoon and the afternoon points often lie on the same straight line, the second term in equation (2) ought to result in the afternoon values of f_0E being greater than the forenoon values for the same $\cos \chi$. In theory,

VOL. 17 1959/6

therefore, the points on the log–log plot ought to lie on an elongated U-shaped curve with its apex near noon. Curves of this type are often observed to occur but they do not lead to serious difficulties because even if separate lines are fitted to the forenoon and afternoon branches, both lines converge near noon and there is seldom any difficulty in deciding on the best value for f_r.

5. The Accuracy of the Daily Values of J_E

5.1. *Erratic errors*

For the hours 0900 to 1500 LMT inclusive, the distribution of the differences between the observed values of f_0E and those defined by the lines used in the determination of f_r was found to be normal and to have a standard deviation (s.d.) of 0·07 Mc/s. Estimates were also made of the differences between the values of f_r used to determine J_E, and the values which would have resulted from the two extreme, but still plausible, positions of the lines. The s.d. of these differences was 0·035 Mc/s but it is considered unlikely that the s.d. of the error in determining f_r can be more than half this figure; this would be the equivalent of an s.d. of approximately 2 per cent in J_E. In Tables 1 and 2, the values which, it is thought, may contain errors exceeding 3 per cent, are shown in parentheses.

It has been pointed out by Denisse and Kundu (1957) and by Minnis and Bazzard (1958) that there is a close correlation between the monthly mean values of the solar radio noise flux, Φ, at λ 10·7 cm and certain E-layer indices. A rough visual check of the daily values of Φ and J_E confirms the existence of this correlation even on a daily basis. It has been noticed, however, that whereas the erratic component in the day to day trend of Φ is very small, the corresponding irregularities in J_E often attain values of ± 4 per cent. The s.d. of the reading errors in the final values of J_E has already been shown to be about 2 per cent and there is little doubt that the erratic fluctuations are partly due to these errors.

5.2. *Seasonal variations*

Minnis and Bazzard (1959) have found that the constants which define the linear relation connecting the mean values of Φ and Ch_E for any given month vary slightly with season. It was shown that these variations originated in Ch_E, the values of which were corrected accordingly to give the values of the index I_E which were published.

Similar corrections ought, in principle, to be applied to the values of J_E given in Tables 1 and 2. This has not been done because, owing to the rapid changes in the mean correction factor from one month to the next, it is not clear how to interpolate a smooth curve which would permit the daily values of J_E to be corrected. For this reason, some caution should be exercised in comparing the absolute values of J_E for different months. As a guide to the magnitude of this effect, the monthly mean corrections appropriate to the period of the IGY are given in Table 3.

5.3. *Geographical variations*

The values of J_E in Tables 1 and 2 are based on ionospheric observations made at Slough during the 8 hr period centred on 1200 UT. It would be desirable

Table 1. Daily values of E-layer index J_E

Date	1957 July	Aug.	Sept.	Oct.	Nov.	Dec.	1958 Jan.	Feb.	Mar.
1	225	272	308	307	364	322	324	290	290
2	280	296	350	328	329	333	(345)	289	290
3	258	291	(305)	324	362	341	342	319	304
4	289	252	297	313	350	(375)	(373)	326	275
5	267	278	303	277	318	(387)	(347)	311	284
6	301	273	331	(302)	310	(330)	331	318	294
7	254	297	—	(312)	314	—	379	303	315
8	234	320	—	337	347	—	(335)	303	305
9	254	285	317	348	292	(369)	336	303	340
10	271	280	(261)	308	292	(357)	329	284	295
11	271	305	339	329	312	299	327	245	306
12	291	294	352	308	296	(343)	(401)	313	296
13	298	283	326	348	339	286	(347)	284	323
14	320	279	339	340	348	(350)	(344)	291	334
15	290	300	352	336	352	333	357	301	—
16	293	321	317	339	323	352	337	309	285
17	297	282	319	344	348	377	414	277	285
18	277	311	368	389	331	314	324	239	255
19	292	(328)	398	397	314	(456)	378	248	(262)
20	307	(301)	378	398	369	284	(382)	255	(297)
21	326	267	421	368	351	(388)	327	248	304
22	305	283	332	(361)	351	(452)	327	240	298
23	294	315	270	(459)	321	(461)	323	255	—
24	304	317	317	(387)	391	(379)	312	273	283
25	336	331	345	(378)	324	402	(353)	261	281
26	330	323	359	(366)	332	(387)	304	274	(317)
27	300	296	333	402	275	416	293	287	314
28	289	(260)	—	355	(290)	351	282	297	339
29	317	294	318	398	312	(423)	(268)		365
30	299	317	266	364	318	367	248		421
31	260	382		336		(352)	293		390

if similar indices could be worked out using data from observatories located approximately 120° east and west of Greenwich. These complementary indices would refer to the hours centred on 0400 and 2000 UT and, in conjunction with the Slough index, would permit the calculation of a composite index which would represent the whole Greenwich day.

If such indices should be worked out for other ionospheric observatories, systematic differences between the different sets of indices may be found to occur owing to the fact that the E-layer ionization for a constant zenith angle and level

VOL. 17 1959/6

Table 2. Daily values of E-layer index J_E

Date	1958 Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1	380	365	294	318	349	347	294	317	314
2	369	353	(263)	318	343	375	321	—	360
3	335	345	(336)	341	341	335	303	300	323
4	293	—	305	319	342	343	303	285	326
5	323	356	(336)	304	311	293	292	274	269
6	328	373	338	(292)	—	324	273	310	309
7	339	344	(292)	317	(342)	307	276	292	337
8	340	311	307	314	360	(305)	279	315	327
9	335	294	291	273	(318)	(324)	295	296	(320)
10	329	281	274	288	(349)	354	326	289	309
11	305	290	(306)	303	(347)	(342)	318	265	329
12	279	277	294	300	(338)	377	311	258	325
13	289	279	293	297	(346)	380	294	231	326
14	281	276	290	256	(337)	(346)	329	268	324
15	311	253	(276)	284	322	(414)	333	274	308
16	322	(286)	262	267	321	(366)	333	285	278
17	323	321	284	268	309	309	376	274	345
18	305	(276)	290	258	(364)	322	345	295	264
19	316	(275)	314	(269)	(366)	291	384	284	(312)
20	318	(271)	296	248	330	314	353	280	342
21	333	294	290	299	366	323	329	272	329
22	321	305	290	279	360	336	323	272	292
23	310	307	327	(312)	338	305	284	297	283
24	(322)	275	(293)	303	(329)	308	298	301	287
25	333	312	320	342	348	310	287	(316)	310
26	345	312	290	333	323	296	—	303	327
27	337	290	299	312	411	293	281	302	270
28	332	323	293	355	344	302	271	271	312
29	383	313	290	349	—	301	291	304	295
30	392	298	300	371	—	346	271	331	322
31		309		(347)	—		284		309

of solar activity increases towards the equator. Data on this latitude effect have been given by ALLEN (1948) and it can be shown that the relation between J_E for Slough and for an observatory at latitude λ is given by

$$\log J_E(\lambda) = \log J_E \text{ (Slough)} + 2{\cdot}18 \times 10^{-3} (51{\cdot}5 - \lambda) \qquad (8)$$

6. CONCLUSIONS

A method has been worked out for deriving a daily index, J_E, of solar ionizing radiation from hourly measurements of the critical frequency, f_0E, of the E-layer.

Precautions have been taken to minimize the errors in identifying and reading f_0E and it is estimated that the residual errors due to these causes lead to fluctuations in J_E having an s.d. of about 2 per cent. Apart from these fluctuations, there is a close similarity in the short-term variations of J_E and the solar radio noise flux at λ 10·7 cm.

Table 3. Monthly mean correction factors to be applied to J_E

Month	Factor	Month	Factor	Month	Factor
1957 July	0·96	1958 Jan.	0·95	1958 Jul.	0·95
Aug.	1·02	Feb.	0·90	Aug.	1·01
Sept.	1·06	Mar.	1·04	Sept.	1·06
Oct.	1·12	Apr.	1·03	Oct.	1·10
Nov.	0·96	May	1·08	Nov.	0·98
Dec.	0·98	June	1·10	Dec.	0·98

Acknowledgements—The work described in the paper was carried out as part of the programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

REFERENCES

ALLEN C. W.	1948	*Terr. Magn. Atmos. Elect.* **53,** 433.
APPLETON E. V. and NAISMITH R.	1939	*Phil. Mag.* **27,** 144.
BEYNON W. J. G. and BROWN G. M.	1956	*Nature, Lond.* **177,** 583.
DSIR	1953	Radio Research Special Report No. 23 H.M.S.O., London.
DENISSE J. F. and KUNDU M. R.	1957	*C. R. Acad. Sci., Paris* **244,** 45.
MINNIS C. M. and BAZZARD G. H.	1958	*Nature, Lond.* **181,** 1796.
MINNIS C. M. and BAZZARD G. H.	1959	*J. Atmosph. Terr. Phys.* **14,** 213.
NICOLET M. and BOSSY L.	1949	*Ann. Géophys.* **5,** 275.
URSI	1958	Information Bulletin No. 111, 24.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 65 to 70. Pergamon Press Ltd. Printed in Northern Ireland

Characteristics of the visual aurorae at Byrd Station, Antarctica, during 1957

D. Payne Hale
USNC–IGY Deepfreeze Two, Expedition to Antarctica

(*Received* 13 *April* 1959)

Abstract—Characteristics of the visual aurorae observed at Byrd Station, Antarctica, located at 80° S latitude, 120° W longitude, during the last half of 1957 are discussed. Salient features of auroral activity variations in time, spatial and motion patterns, colour frequency, vortices, swirling rays, special forms and other aspects are delineated. The most active period over the season during the Greenwich sidereal day is from 0200 to 0600 GCT. Aurorae of asymmetric intensity relative to the geomagnetic meridian were observed, with the brightest portion about three times more likely to occur east rather than west of the meridian. Off-meridian aurorae and bands of long duration were observed. The direction of motion of features was overwhelmingly from west to east; and forms were much more likely to drift northward than southward. A possible method for measuring ray-lengths employing but a single point of observation is sketched.

OL.
17
9/60

Introduction

The observations discussed in this paper are primarily those of a half season at Byrd Station, Antarctica, located at 80° S latitude, 120° W longitude. Base construction and laboratory set up prevented operation of a systematic aurora programme until well into the Antarctic winter. Observations were made using an alidade and the well-known Cornell aurora diagram and report cards (IBM), supplemented by on the spot notes on the back of the diagram card. The period of regular observation, on the hour every hour for most of the season's duration, was from 30 June 1957 to well beyond the last aurora, occurring on 24 September 1957.

Time Variations

An aurora was observed at Byrd Station on 6 April; this was the first noticed. Aurorae could have occurred prior to this date since efficient watch for them could not be made at that time. Fairly active, whitish, aurorae occurred on 7 April. Bright green aurorae were seen on 21 April, followed by a red corona in the early hours of 24 April. Here and throughout this paper, all dates and other times are GCT.

Vigorous displays were seen on the following dates, each day having several hours of aurorae, and 30 August, over 10 hr of almost continuously active aurora.

30 June (special world interval)	18 August
15 July	19 August
19 July	30 August (special world interval)
10 August	9 September
15 August	30 September

The most active day of the season was either 30 June or 30 August—both SWI's. 30 June had at least six separate displays of red aurorae of medium or brighter intensity; and 30 August had four very bright aurorae and two bright to very

bright, with many others of lesser intensity, all occurring within a 45 min period (0200 to 0300 GCT) and belonging to the same display.

An analysis of daily variation in auroral frequency weighted for intensity (multiplicative factors as follows: faint, unity; medium, two; bright, three; very bright, four; with in-between intensities upgraded half a rank) indicates the diurnal pattern below:

0200 to 0600 GCT, the most active hours throughout so long as the night condition maintains.
2200 to 0200 GCT and 0600 to 1000 GCT, periods of fair activity.
1000 to 2200, the dullest period; local noon is 2000 (GCT).

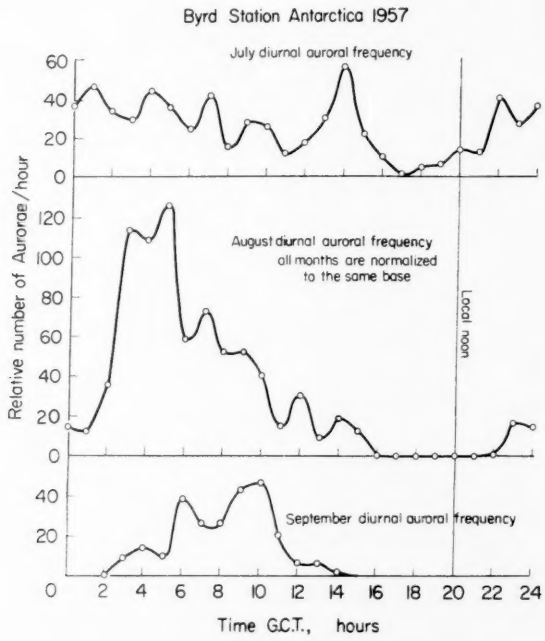


Fig. 1.

Twilight, toward the end of the season, of course drastically affects the activities of these periods, which it is to be understood are characteristic of the predominant night condition. As the season neared termination hours of peak activity shifted progressively later in the "day," with 0600 to 1100 (GCT) being the period of activity during September. Local midnight is 0800 (GCT).

Pulsations of periods ranging from a few minutes to a few seconds were observed; the median period was around 30 sec. Pulsations seemed to characterize the beginning and ending of strong displays; however, some may have occurred during high activity and passed unnoticed due to the observer's attention being elsewhere.

VOL. 17 1959/6

Spatial Patterns

A surprising tendency for aurorae to be brighter in off-meridian portions was noticed. It cannot be discounted as a simple Van Rhijn effect in that in individual cases it is asymmetric about the magnetic meridian. A related feature is that about two-thirds of the aurorae of this type had their brightest part east of the meridian, with the remaining third west.

Occasionally "half-forms" were observed, so named since such forms appear on only one side of the meridian, and are normal in other respects. Nineteen such cases were noted, of which twelve were east of the meridian and the remainder west. This is in line with the observations of the paragraph immediately above; however, both sets of observations are enigmatic in that they tend to give special significance to the observer's meridian. There does not seem to be any relation between the time of appearance of these asymmetric intensities, and the side of the meridian upon which they manifest.

It was frequently noticed that there appeared to be an area of higher activity, in both auroral frequency and activity, east of Byrd Station. Assuming the average height of Byrd aurorae to be about 60 miles, this locale is between 150 to 200 miles east of Byrd Station.

About 200 miles to the geomagnetic north there often appeared during moderate to strong displays, what was named "the setting band." A homogeneous band would appear in one of the zones* LN, N4 or N3, and remain there for hours, slowly shifting its position within these three zones. This band was observed on seventeen occasions, and it had an average duration of about 4 hr. On 30 August it lasted for 11 hr. In the sense of perhaps being the most probable"parallel of geomagnetic latitude" for particle precipitation, the setting band may be the aurora zone itself.

Off-meridian aurorae were frequent; by off-meridian aurorae is meant symmetric aurorae whose symmetry axes make appreciable angles with the geomagnetic meridian. For example, by northern off-meridian aurora is meant an aurora whose axis of symmetry runs true north from the observer.† Off-meridian eastern aurorae seem to be more frequent during mid-season, with off-meridian western forms predominating toward the end of the season. Western forms were more of the nature of ray bundles of short duration (in which case azimuth can be used directly without considering a symmetry axis), while eastern aurorae appeared more frequently in stable forms such as arcs and bands. On the basis of number the western forms exceed the eastern; however, a time–intensity integration would probably favour the eastern over the western. There were five instances of aurorae parallel (i.e. symmetry axis perpendicular to) with the geomagnetic meridian.

Patterns of Motion

Undoubtedly many motions were not observed due to the observer's full attention being required elsewhere. However, of those which were noted, the following patterns were evident.

* Zone designation is that of the Cornell aurora diagram card.

† This is recognized to be not precise in that the symmetry axis of an off-meridian aurora generally does not intersect the observer's meridian at the observer's position.

(1) An aurora was three times more likely to drift from west to east than vice versa.

(2) An aurora was about three times more likely to drift northward than southward. One conjectures that aurorae on the average tend to drift toward the auroral zone.

In regard to observations (1) and (2) there was definitely no simple correlation between the direction of drift and local geomagnetic time, i.e. no effect such as that stated for the northern hemisphere of westward drift before magnetic midnight and eastward drift after MMN.

(3) The direction of flow taken by rays and undulations along normally oriented arcs and bands is predominantly (about 8 to 1) from geomagnetic west to geomagnetic east. On 17 June at 1143 GCT, rays were observed apparently rippling in *both* directions simultaneously.

(4) Swirling rays, that is ray bundles which advance with an epicycloidal motion, were observed on 29 July at 0332, 10 August 0227, 18 August 0328 and 9 September 0230—all GCT. Dates and times are given because these were aurorae of unusual vigour. I was unable to make certain of the rotation sense. On 20 July 0238 GCT, an unrayed patch was observed swirling clockwise. See further discussion under "special forms and features."

Colour and Colour–Time Correlations

Coloured forms were observed in the following order of decreasing frequency of occurrence, and possible time correlations are suggested in the remarks.

White, throughout season.

Green or yellow green, common except toward season's end.

Red lower border (type *B* aurorae due to 1 P.G. of N_2), from 2100 to 0600 GCT; frequency may follow that of aurorae in general.

Red (type *A* aurorae due to oxygen), usually between 0200 and 0400 GCT.

Yellow, quite distinct from the yellow green, there were eleven such aurorae in three days, all occurring between 0200 and 0300 GCT and near the end of the season (24 September).

Violet grey, between 0900 and 1200 GCT, these aurorae were all seen on 13 and 14 September.

Red upper border, observed only once, a medium rayed band in zone S2 on 24 August at 0408 GCT.

Blue, observed once, a medium corona in the zenith at 1036 on 14 September.

About one-third of the "days" from 30 June to the season's end on 24 September had at least one coloured aurora of a colour rarer than green. Approximately equal numbers of type *A* (uniformly red) and type *B* (red lower border) aurorae were observed, and both types occurred throughout the season.

Intensities

During the half season of regular hourly observations, eight very bright, and twelve bright to very bright aurorae were observed. Five of the very bright aurorae had red lower borders.

VOL. 17 1959/6

Discussions with four observers, three of whom actually participated in the recent Antarctic IGY Aurora Programme, and all of whom have seen both Arctic and Antarctic aurorae, brought forth the unanimous opinion that Arctic aurorae are more intense than Antarctic. One realizes of course that their comparisons are not all based on conjugate geomagnetic locations, and were made at different times; however, this does not discount the fact, that two of these observers experienced their Antarctic aurorae at a station very near the Antarctic zone, and at a time near the peak of the 11-year cycle. Such circumstances would on the contrary tend to make the Antarctic displays seem brighter relative to the Arctic aurorae, and certainly would not favour the Arctic ones. One of these two observers was at Big Delta, Alaska, which is near the Arctic zone; however, the other experienced his Arctic aurorae at Goose Bay, Labrador which is well off the Arctic zone.

Intense forms (i.e. bright to very bright, and very bright) were usually yellow green accompanied by tinges of red; however, one very bright and one bright to very bright aurora were of the rare yellow colour.

Violet aurorae in all cases appeared to be of medium or less intensity—an effect that may be attributable to the decreased sensitivity of the human eye at these wavelengths. About 80 per cent of the red aurorae were of medium or less intensity. Of the eleven yellow aurorae all were bright but two, one being bright to very bright and the other very bright.

OL.
7
9/60

Vortices, Special Forms and Features

Six vortices were observed after 30 June. Five of these were counter-clockwise (CCW) proceeding inward. A CCW vortex was also observed on 17 June. Thus CCW vortices predominated over CW vortices six to one in this small sample.

Twenty-five special forms as classified in column 18 of the IBM Cornell aurora report card were observed. Usually these were not oriented such that column 18 could be used.

A characteristic following vigorous displays was the appearance of large faint and faint to medium diffuse surfaces over large areas of the sky. These frequently persisted for 1 hr or more, and on one occasion they endured for 4 hr. They are quiescent, and often of ill-defined boundaries. Sometimes this condition over small areas seemed to precede more active phases; however, precursor forms for active phases were usually faint to medium flickering rays in the north-west.

Eleven aurorae were of such vigour that they were classified as flaming aurora or near flaming aurora. Of these, three involved violent but simple ray or undulatory motion (e.g. from west to east), or the brief appearance of many short, rapidly moving rays in a small area. The remainder all had features, usually rays but in one case a diffuse surface, which revolved around the magnetic zenith (four cases), or contained ray bundles which advanced with an epicycloidal motion—a phenomenon which was named *swirling rays*. Bundles of swirling rays usually contained both red and green rays, and in two cases the red rays definitely advanced ahead of the green, sweeping through an area about $\frac{1}{2}$ sec before the appearance of the green rays. So sudden, spectacular and brief is this phenomena of swirling rays, that the sense of rotation could not be definitely determined. Revolving rays

definitely turned CCW in all cases; the diffuse surface, mentioned above, rotated CW and seemed to contain the magnetic zenith.

In three observations arcs were seen to break up into regularly spaced segments which were called "kinks." These were noticed on 17 July, 0000 to 0100, and 19 July, 0200 to 0300 and 2200 to 2300. Each segment was hooked at one end, and at the same end for each set of segments. The segments were brighter at the hook end, and the angular "wavelength" (hook to hook), while fairly uniform in each case, varied from 12 to about 28 degrees.

Some ray-lengths were measured by a graphical method based upon two assumptions:

(1) That the base of the ray is at the usual height (100 km). If it is higher than 100 km, this assumption will result in the measured length being less than the true length and a lower height for the top of the ray than its actual height.

(2) That the ray lies along a line of magnetic intensity whose direction in the portion of interest differs negligibly from that of the surface magnetic field at the point of observation.

Crude corrections for the angle between the vertical plane, containing the observer and the base of the ray, and the ray were made. The dip at Byrd Station is about minus 75 degrees, and the declination is 69 degrees E. This method indicated seven rays to be over 400 miles long, with three of these substantially exceeding 500 miles (800 km). Such a procedure may be immediately criticized by citing Byrd Station auroral data proving that the magnetic zenith differs from the average dip, sometimes as much as 3 degrees; however, in one case of very long rays, the magnetic zenith was simultaneously well defined by a corona. The application of this method and its criticism are planned to be the subject of a future paper.

VOL.
17
1959/6

Acknowledgement—The writer here expresses his appreciation to the Air Force Cambridge Research Center, to N. J. OLIVER, Aurora Antarctic Project Leader for the U.S. for the International Geophysical Year, and to the Arctic Institute of North America—all of whom made possible his Antarctic duty. The great help of G. TONEY, Byrd Station Scientific Leader, in the analysis of the data upon which this paper is based, is gratefully acknowledged.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 71 to 81. Pergamon Press Ltd. Printed in Northern Ireland

The determination of the electron distribution in the upper ionosphere from satellite Doppler observations

F. H. Hibberd and J. A. Thomas
University of New England, Armidale, N.S.W., Australia and University of Queensland, Brisbane, Australia

Abstract—It is shown how the electron distribution in the upper regions of the ionosphere may be determined from simultaneous measurements of the Doppler shifts of radio signals of two differen frequencies received from an artificial satellite. The method is illustrated in detail by computations made on the 20 and 40 Mc/s signals from *Sputnik* I.

OL.
17
9/60

1. Introduction

This paper describes how the distribution of electron density with height, in the upper ionosphere, may be deduced from simultaneous measurements of the Doppler shifts in the frequencies of two radio waves emitted from an artificial satellite.

If there were no ionosphere the Doppler shift in the frequency of a radio signal emitted from an artificial satellite, and received on the ground, would be determined solely by the component of the velocity of the satellite in the direction of the straight line between the satellite and the receiver. The presence of the ionosphere modifies this Doppler shift, first because refraction causes the received ray to leave the satellite in a direction different from that of the straight-line path, and secondly because the ionosphere adjacent to the satellite makes the phase velocity there differ from the free-space value. In principle, one could compute the "free-space" Doppler shift from a knowledge of the orbit elements of the satellite and then attempt to find a model ionosphere that would account for the difference between the free-space Doppler shift and the observed shift. Such a procedure would require a knowledge of the orbit elements to a very much higher degree of accuracy than one is ever likely to possess. However, when signals of two different frequencies can be observed simultaneously, it is possible to extract the ionospheric effect from a comparison of the simultaneous variations of their frequencies with time. In this way one eliminates the need to know the orbit data and the absolute frequency of the emitter to an impossibly high order of precision.

2. Effect of the Ionosphere on the Doppler Shifts

The Doppler shift Δf of the frequency of a radio signal received from a satellite travelling in the ionosphere has been shown (Hibberd, 1958) to be

$$\Delta f = \frac{f_0 V \mu_s}{c} (\cos \alpha \sin \psi_s - \cos \beta \cos \psi_s) \tag{1}$$

where f_0 is the frequency of the emitter, V is the velocity of the satellite relative to the observer, and α and β are the direction cosines of V relative to the X- and Y-axes, respectively, as shown in Fig. 1. The refractive index at the satellite is denoted by μ_s, and ψ_s is the angle which the observed ray, as it leaves the satellite,

makes with the inward radial direction *SO*. The effect of the earth's magnetic field is small for the frequencies considered and it is neglected in this work.

We shall be concerned later with the analysis of the Doppler shifts observed with *Sputnik* I, for which the radiated frequencies were approximately 20 and 40 Mc/s, respectively. Quantities associated with the 20 and 40 Mc/s signals are denoted by $'$ and $''$, respectively, and the subscript zero denotes the frequency of

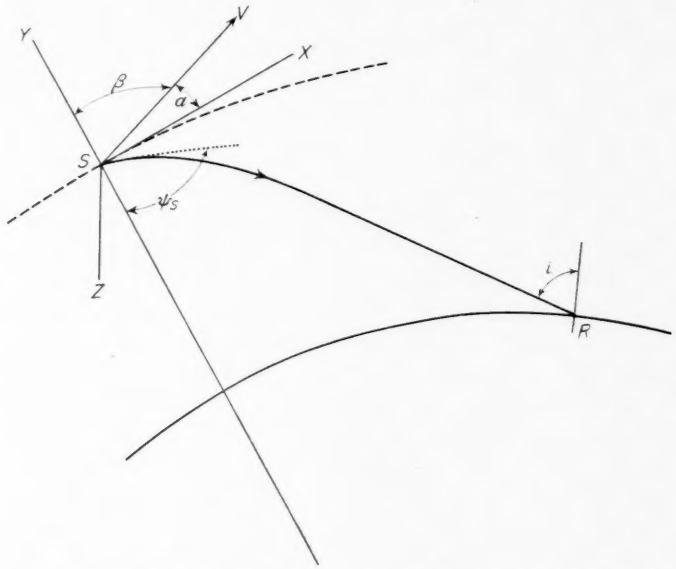


Fig. 1. *S* represents the satellite travelling in the ionosphere with velocity *V*. The receiver is on the ground at *R*. The centre of the earth, *O*, lies along *YS* and is not shown.

a signal as it is emitted from the transmitter. From equation (1) we can write for the frequency of the 20 Mc/s signal as it is received:

$$f' = f_0' + \frac{f_0' V \mu_s'}{c} (\cos \alpha \sin \psi_s' - \cos \beta \cos \psi_s')$$

A similar expression can be written for the 40 Mc/s frequency f''. The difference $2f' - f''$ is then easily shown to be

$$\begin{aligned} 2f' - f'' = D &+ \frac{DV\mu_s}{c} (\cos \alpha \sin \psi_s' - \cos \beta \cos \psi_s') \\ &+ \frac{Vf_0''}{c} [\cos\alpha (\mu_s' \sin \psi_s' - \mu_s'' \sin \psi_s'') \\ &- \cos \beta (\mu_s' \cos \psi_s' - \mu_s'' \cos \psi_s'')] \end{aligned} \qquad (2)$$

where $D = 2f_0' - f_0''$, that is, D is a constant.

The second term in (2) represents a small fluctuation of approximate amplitude DV/c and may be neglected. (For *Sputnik* I $D \sim 7$ kc/s and $V \sim 7$ km/sec, so that $DV/c \sim 0{\cdot}2$ c/s.) The effect of the ionosphere on the difference $2f' - f''$ is

therefore represented entirely by the third term in (2), so that $2f' - f''$ varies with time in the same manner as the expression

$$(Vf_0''/c)[\cos\alpha(\mu_s' \sin\psi_s' - \mu_s'' \sin\psi_s'') - \cos\beta(\mu_s' \cos\psi_s' - \mu_s'' \cos\psi_s'')] \quad (3)$$

In the absence of the ionosphere (or when the satellite is effectively below the ionosphere) $\mu_s' = \mu_s'' = 1$ and $\psi_s' = \psi_s''$. Expression (3) is then equal to zero and, according to (2), a graph of $2f' - f''$ against time should be a straight line of constant frequency D on which is superimposed a fluctuation whose amplitude does not exceed about 0·2 c/s.

3. Outline of Method of Obtaining Electron Distribution

The method consists of finding, by trial and error, the distribution of electron density which, when combined with the motion of the satellite, gives the same variation of $2f' - f''$ with time as that observed experimentally.

Orbit elements deduced from radio observations are sufficiently accurate for the present purpose. A number of convenient instants are selected during the interval for which the Doppler measurements are available, and for each of these instants the position of the satellite and the quantities V, α and β are calculated.

The electron distribution from the base of the ionosphere up to the level of the F-maximum, deduced from conventional ionospheric sounding measurements, is combined with an assumed distribution from the F-maximum up to the maximum height of the satellite to give a model for the whole ionosphere. For this model ionosphere one can calculate the values of μ_s at the height of the satellite at each of the selected instants, for each frequency. The values of ψ_s are obtained by tracing the rays through this ionosphere from the satellite to the Doppler measuring station on the ground. One can then evaluate $2f' - f''$ at each instant and compare the resulting graph of $2f' - f''$ against time with that observed experimentally. By repeating this process for various assumed distributions above the F-maximum the distribution that gives the best agreement with observation can be found.

Most of the remainder of this paper is devoted to an account of an application of the method to some Doppler measurements made on the 20 and 40 Mc/s signals from *Sputnik* I. The ray tracing was done by manual computation and for this reason a simple uniform ionosphere from an arbitrary level upwards was chosen. Such a model ignores the existence of the maximum in the distribution. The consequences of this are examined. For more realistic models the computations should be carried out with an automatic computer. The main purpose of the following is to illustrate the method and demonstrate that it is practicable.

4. The Doppler Observations

The most extensive and accurate data available to us for both frequencies simultaneously were those obtained by the British Post Office Frequency Measuring Stations at Baldock and Banbury. In these, the frequencies of each signal had been measured to within about 1 c/s at intervals of about 10 sec. A preliminary examination of the data was made by plotting each frequency against time on separate sheets of transparent paper with the frequency scale for the 20 Mc/s

values double that for the 40 Mc/s values. Since the ionospheric effect is small, these curves should be almost identical for any particular transit. This was checked by superimposing the curves, with the corresponding times coincident and sliding the graphs along the frequency axis to see if coincidence of the curves could be obtained. On some records effects of multipath and anomalous propagation, and also small frequency fluctuations that are perhaps associated with scintillations, were detected, especially when the satellite was distant from the receiving point. Only those records that were free from defects of the above type were retained for further examination for the present purposes.

Before taking the difference $2f' - f''$ for these records a correction had to be applied because the 20 and 40 Mc/s frequencies had not been measured in an identical manner. One frequency had been measured "instantaneously"; the other had been measured by integrating the number of cycles over 1 sec. In order to obtain "simultaneous" values it was necessary to displace the recorded times of measurement of these latter frequencies by $\frac{1}{2}$ sec before taking the difference. Also, a small amount of smoothing was cautiously applied to each of the frequency–time curves by examining the first and second differences of successive values. Differences $2f' - f''$ were then evaluated at 10 sec intervals. Further, when data were available, the resulting graphs of $2f' - f''$ against time, obtained from Baldock and from Banbury for the same transit, were compared to see whether they were similar, allowing for the small expected difference between them.

It was mentioned in Section 2 that if the ionosphere were absent the graph of $2f' - f''$ against time should be a straight line at constant frequency. We find for the records obtained when the satellite was in perigee near the observing point, and so was effectively below all of the night-time ionization, that the graph is practically a straight line.

The effect when the satellite is travelling in the ionosphere in the day-time should be reasonably large, but unfortunately no satisfactory day-time observations were available. Also, we have not considered records near sunrise or sunset because of the complications associated with horizontal gradients of the ionization. The ionospheric effect at night is extremely small and, because of this, only a few sets of data possessed the required precision (of the order of 1 part in 10^8) to warrant further detailed examination. One such set was for the transit of 0340 hours on 17 October 1957. The experimental curve of $2f' - f''$ against time, observed at Baldock, is reproduced as curve (a) of Fig. 6. Only portion of the corresponding curve for Banbury was available but it was in good agreement with that for Baldock.

5. Orbit Data and Computations

The orbit elements used are those published by the Mullard Radio Astronomy Observatory (1957). The values of the elements differ considerably from those obtained by the Royal Aircraft Establishment (1957) but there are at present no criteria by which one can judge which are the better values. Both orbits are admittedly provisional but accurate orbit data, such as would be obtained from optical or radar tracking, are unfortunately not yet available. The following data have been used in the computations:

Time of Cambridge Meridian Transit: 17 October 1957, 0342 hours, 16 sec.

Latitude of Cambridge Meridian Transit $= 55°30'$.

VOL.
17
1959/6

Distance and bearing of Baldock from Cambridge: 30 km; 141°47′ W of N.
Orbital period T at 17 October, 0342 hours = 5745·3 sec.
Semi-major axis at 17 October, 0342 hours = 6929·4 km.
Eccentricity of orbit $e = 0{\cdot}053$.
Inclination of orbit = 64°40′.
Latitude of Perigee = 36°N.

We introduce the following definitions and formulae of planetary motion (see, e.g., SMART, 1949):

Mean angular motion $n = 2\pi/T$.
Instant of perigee τ.
Mean anomaly M.
Eccentric anomaly E.
True anomaly v.
Radius vector of satellite, from centre of earth r.
Orbital velocity V.*

$$\dot{r} = \text{component of } V \text{ along } r$$

$$M = n(t - \tau) = E - e \sin E \tag{4}$$

$$\tan \frac{V}{2} = \left(\frac{1+e}{1-e}\right)^{\frac{1}{2}} \tan \frac{E}{2} \tag{5}$$

$$r = a(1 - e \cos E) \tag{6}$$

$$V = 2n^2a^3/r - n^2a^2 \tag{7}$$

$$\dot{r} = nae(1 - e^2)^{1/2} \sin v \tag{8}$$

In calculating the velocity and position of the satellite at a number of instants it is convenient to work in terms of the true anomaly v and to measure times from the instant of perigee rather than from the instant of closest approach to the Doppler measuring station at Baldock. The calculation of the instant when the satellite is in perigee is assisted by reference to Fig. 2. The figure represents a hemisphere, with centre at O, whose radius is equal to that of the earth. $ABCD$ represents the projection on the surface of the earth of the northern portion of the orbit that crosses the equatorial plane $AODFE$ at A and D. $GBCH$ is a parallel of latitude λ, and P is the north pole. It is seen from the figure that

$$\sin AB = \frac{\sin \lambda}{\sin i} = \sin AC \tag{9}$$

For the particular transit considered, the satellite was travelling southward and rising when near Baldock. Thus the angular distance from its last equatorial crossing to its crossing of latitude λ, when travelling southward, was AC.

We take as our initial time and position data those determined by combined interferometer and Doppler observations at Cambridge. The satellite crossed the Cambridge meridian at 0342 hours 16 sec at a ground latitude of 55°30′. Equation (9) then gives the distance of the Cambridge meridian crossing from the previous equatorial passage as 114°15′. The latitude of perigee was 36°N. Using this in equation (9) gives the distance of perigee from the equatorial passage as 40°34′. The distance of the Cambridge meridian crossing from perigee was therefore 114°15′ − 40°34′ = 73°41′. This is the true anomaly v at the Cambridge meridian crossing. By means of equations (4) and (5) we then obtain the result that the satellite was in perigee 1084 sec before it crossed the Cambridge meridian; that is, the instant when it was in perigee was 0325 hours, 12 sec.

Further, we know that the projection of the orbit cuts the Cambridge meridian 360 km N of Cambridge, that the distance and bearing of Baldock from Cambridge are 30 km, 141°47′W of

* Throughout the following the velocity of the satellite relative to the observer has been taken to be the same as the orbital velocity, and the rotation of the earth has been neglected. The velocity of the observing point at latitude 52°, due to rotation, is about 0·3 km/sec whereas the orbital velocity is about 7 km/sec. In precise calculations the effects of rotation should be taken into account but their inclusion in the present work is not warranted.

N, and that the inclination of the orbit is 64°40′. From these we find that the distance from the Cambridge meridian crossing to where the orbit was closest to Baldock was 2°09′. The true anomaly at the point of closest approach to Baldock was therefore 73°41′ + 2°09′ = 75°50′. Using this result in equations (4) and (5) shows that the instant of closest approach to Baldock occurred 1117 sec after perigee; that is, that the time of closest approach to Baldock was 0342 hours 49 sec. As a check, this may be compared with the time of occurrence of maximum slope in the Baldock Doppler curves of 0342 hours 46 sec ± 4 sec for the 20 Mc/s and 0342 hours 50 sec ± 4 sec for the 40 Mc/s. It is thus seen that the presence of the ionosphere does not cause

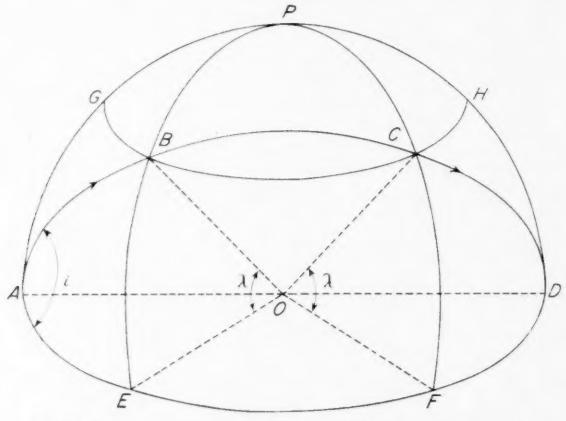


Fig. 2. The curve $ABCD$ represents the projection of the northern portion of the orbit on the surface of the earth.

any appreciable time displacement of the Doppler frequency–time curve and that the instant of closest approach of the satellite to the observer may be taken as the instant of occurrence of maximum slope in the Doppler curve.

We now select a number of convenient instants t, within an interval ±4 min about the time of closest approach (c.a.) 0342 hours 49 sec and, taking $t-\tau$ for each, compute in turn r, V and $\dot{r}$ by means of equations (4) to (8).

The direction cosine cos β at each instant is obtained from

$$\cos\beta = \dot{r}/v \tag{10}$$

The direction cosine cos α is obtained from

$$\cos\alpha = \sin\beta\cos\alpha' \tag{11}$$

where

$$\sin\alpha' = \frac{\sin(\text{ang.}^r \text{ dist. of c.a.})}{\sin\gamma} \tag{12}$$

The derivation of equation (12) is illustrated by reference to Fig. 3; α' is the angle on the ground between the projection AC of the orbit and the projection AB of the ray path from satellite to observer; Δv is the angular distance between the position of the satellite at the instant considered and its position when it is at closest approach to the observer; γ is the angular distance between the satellite and the observer at the instant considered. The angle γ is calculated from

$$\cos\gamma = \cos(\text{ang.}^r \text{ dist. of c.a.})\cos\Delta v \tag{13}$$

6. Ionospheric Data

Hourly records of h′(f) curves for 17 October 1957 were available from the Radio Research Station at Slough. These were reduced to $N(h)$ curves by the computing group of the Radio

VOL. 17 1959/6

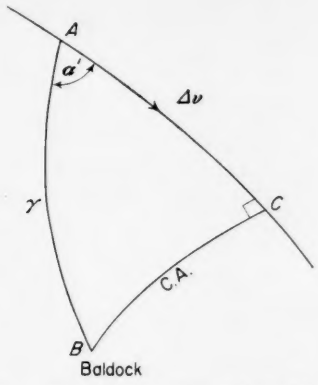


Fig. 3. The figure shows the projection Δv of part of the orbit on the surface of the earth when the satellite is near its point of closest approach (C.A.) to the receiving station at Baldock.

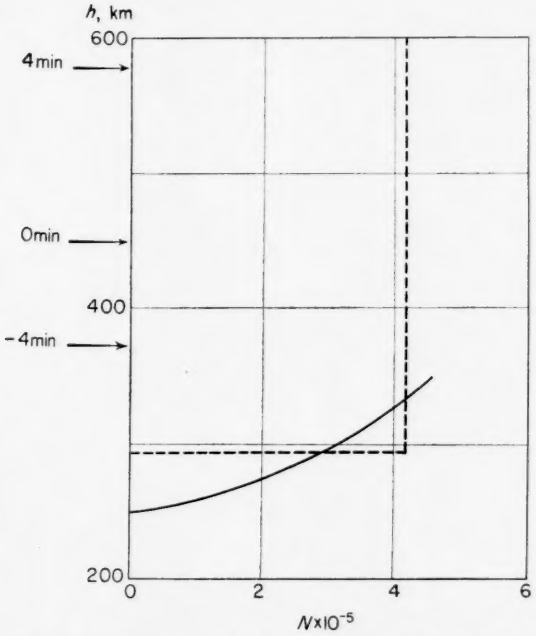


Fig. 4. The full line represents the N–h curve existing at the time of the Doppler observations. The dotted curve is the N–h model used in the computations. The arrows on the left indicate the height of the satellite at the instant of closest approach to Baldock and at 4 min before and after this, respectively.

Section of the Cavendish Laboratory. Using these, especially that for 0400 hours, a uniform model ionosphere was chosen with the following characteristics:

height h_0 of lower boundary $= 295$ km;
electron density $N = 4{\cdot}2 \times 10^5$ cm $^{-3}$.

For this ionosphere, $\mu' = 0{\cdot}9580$, $\mu'' = 0{\cdot}9895$. Fig. 4 shows the actual $N(h)$ curve for 0400 hours,

the curve for the model, and the heights of the satellite at -4 min, 0 and $+4$ min, respectively, measured from time of closest approach.

The crude model with uniform electron density was chosen solely to simplify the ray tracing necessary to compute the angle ψ_s at which the observed ray leaves the satellite.

7. Ray Tracing

Fig. 5 is drawn in the plane containing the ray SCR from satellite S to observer R. The point C is at the base of the ionosphere, O represents the centre of the earth, and the radius of the

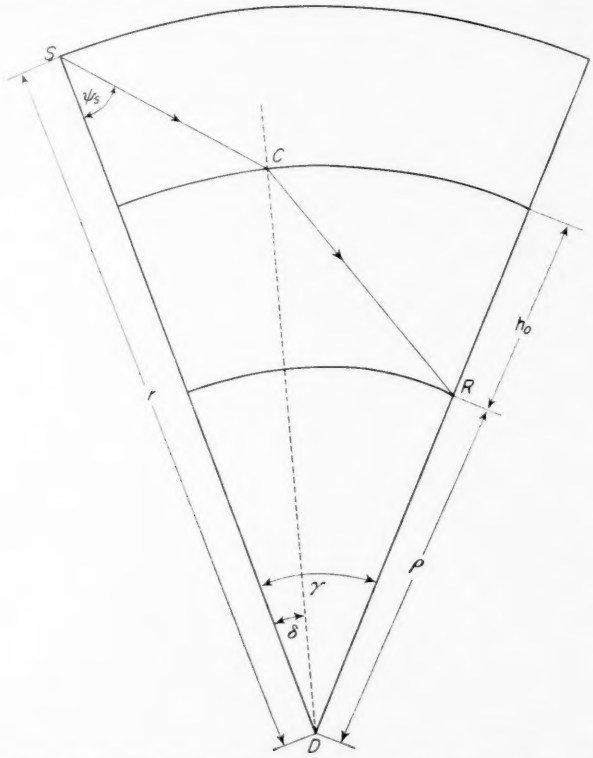


Fig. 5. This figure illustrates the geometry used in determining the angle ψ_s. SCR is the ray from satellite S to the receiver R on the ground. C is at the base of the ionosphere. O represents the centre of the earth.

earth is denoted by ρ. The optical path length P of the ray SCR, for a given value of γ, must satisfy the condition $\partial P/\partial\delta = 0$. This condition yields the equation

$$\mu r \sin\delta[r^2 + (\rho + h_0)^2 - 2r(\rho + h_0)\cos\delta]^{-1/2} + r\sin(\gamma - \delta)[(\rho + h_0)^2 + \rho^2 - 2R(\rho + h_0)\cos(\gamma - \delta)]^{-1/2} = 0 \tag{14}$$

For each value of γ, corresponding to each of the selected instants in the satellite's travel, equation (14) was solved graphically for the angle δ, using the refractive indices for the 20 and 40 Mc/s signals in turn.

From Fig. 5 it is seen that

$$(\rho + h_0)\sin(\psi_s + \delta) = r\sin\psi_s$$

therefore

$$\tan\psi_s = \frac{\sin\delta}{r/(\rho + h_0) - \cos\delta} \tag{15}$$

which allows ψ_s to be calculated from δ.

VOL. 17 1959/6

8. Comparison of Computed and Observed Differences

Using the values of V, α and β computed from the orbit data, ψ_s obtained from equation (15), the nominal value of 40 Mc/s for f_0'' and the values of μ_s' and μ_s'' appropriate to the model ionosphere, the value of expression (3) was computed at each of the selected instants. The computed points are shown in Fig. 6 (curve b) with an arbitrary choice of zero on the frequency axis, for comparison with the observed curve (a) of $2f' - f''$.

The two curves in Fig. 6 are similar in shape, indicating that the procedure so far followed is essentially correct. However, the magnitude of the variation in

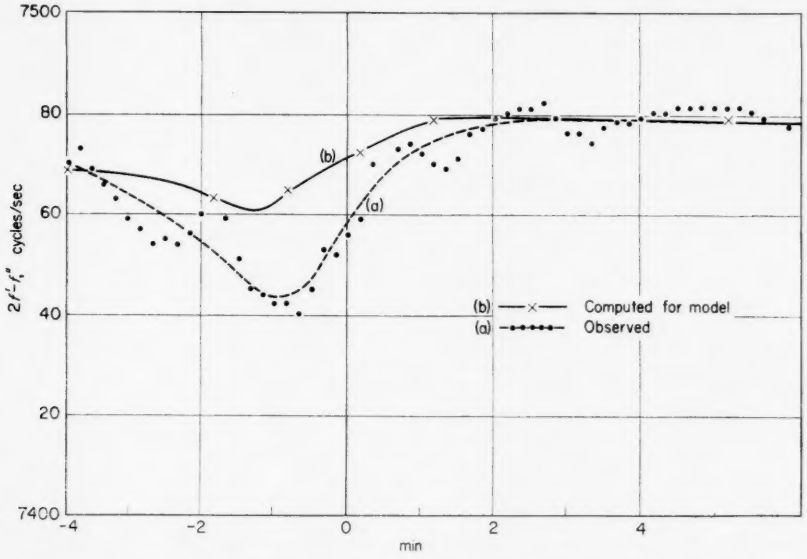


Fig. 6. The dashed line (a) is drawn through the dotted points which represent the differences $2f' - f''$ obtained from the observed 20 and 40 Mc/s signals. The full line (b) is drawn through the crosses which represent the differences $2f' - f''$ corresponding to the model ionosphere shown in Fig. 4.

curve (b) is less than that in curve (a) by a factor of about 2, and the minimum in (b) occurs some 20 sec earlier than that in (a).

9. Effect of the Maximum in $N(h)$

It will now be shown that the differences between curves (a) and (b) in Fig. 6 can be accounted for by the neglect, in the uniform ionospheric model, of the existence of a maximum in the vertical distribution of electron density. Fig. 7(a) represents the ray paths from satellite S to observer R for a uniform ionosphere, for frequencies f' and f''. It is seen that $\psi_s'' < \psi_s'$. Fig. 7(b) represents the ray paths through an ionosphere consisting of a uniform slab of ionization with free space above and below the slab. (This corresponds roughly to transmission from a satellite that is above the level of maximum ionization). Here the rays for the two frequencies intersect within the slab, and consequently $\psi_s'' > \psi_s'$.

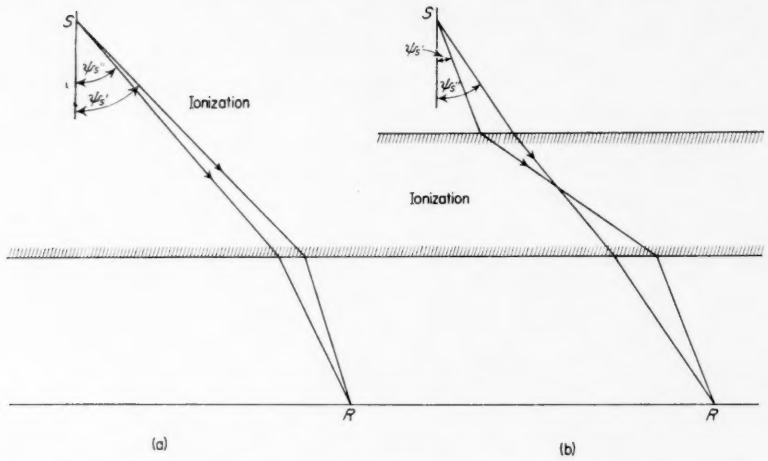


Fig. 7. The figures show the respective ray paths for 20 and 40 Mc/s signals for two different types of ionospheric model. In (a) the electron density is uniform and extends indefinitely upwards from the base the ionosphere. In (b) the ionosphere consists of a uniform slab of ionization, with the satellite above the upper boundary of the slab. This figure corresponds roughly to a satellite travelling in the ionosphere above the height of maximum ionization. It is seen that in (a) $\psi_s'' < \psi_s'$ whereas in (b) $\psi_s'' > \psi_s'$.

We now refer to expression (3), which determines the magnitude of the variation of $2f' - f''$. For the present qualitative discussion we neglect the term multiplied by $\cos\beta$, since $\cos\beta \ll \cos\alpha$, and consider the dominant part

$$(Vf_0''/c)\cos\alpha(\mu_s'\sin\psi_s' - \mu_s''\sin\psi_s'')$$

For $f_0' < f_0''$ it is known that

$$\mu_s' < \mu_s''$$

For a uniform ionosphere (Fig. 7a)

$$\sin\psi_s' < \sin\psi_s''$$

and for an ionosphere with a maximum in $N(h)$ (Fig. 7b)

$$\sin\psi_s' > \sin\psi_s''$$

Thus the magnitude of the difference

$$\mu_s'\sin\psi_s' - \mu_s''\sin\psi_s''$$

is always greater for transmission through a layer with a maximum in N than through a uniform ionosphere. To change from a model such as that in Fig. 7(a) to one such as that in Fig. 7(b), while retaining the same total electron content in each model, can easily increase the magnitude of the difference $\mu_s'\sin\psi_s' - \mu_s''\sin\psi_s''$ by a factor of several times. Thus the difference between the computed and observed curves in Fig. 6 could easily be removed by using a model ionosphere with a maximum in the electron density below the satellite.

The displacement of the time of occurrence of the minima of the curves in

VOL. 17 1959/6

Fig. 6, from the time of closest approach, is due to the radial component of the velocity of the satellite. This effect is represented in (3) by the quantity

$$\cos\beta(\mu_s' \cos\psi_s' - \mu_s'' \cos\psi_s'')$$

It is clear from Fig. 7 that the magnitude of this quantity is greater with a uniform ionosphere than with an ionospheric layer.

We therefore conclude that the observations on the Doppler shifts of signals from *Sputnik* I are consistent with the known vertical distribution of electron density below the level of maximum ionization.

10. Suggested Procedure for Determining $N(h)$ above h_{max}

When the electron density does not vary monotonically with height between observer and satellite it is not possible to assign unique values of μ_s and ψ_s consistent with the observed effect. If, however, $N(h)$ is known up to h_{max} and N is assumed to decrease monotonically with h above h_{max}, there will be a unique $N(h)$ curve above h_{max} consistent with the observed effect. The procedure to determine $N(h)$ above h_{max} may be summarized as follows:

(i) From simultaneous measurements of the two frequencies, obtain $2f' - f''$ as a function of time, near the instant of closest approach.

(ii) Compute the parameters V, α, β of the orbital motion at a number of instants near closest approach.

(iii) Obtain the $N(h)$ distribution up to h_{max} from ordinary ionosonde observations.

(iv) Adopt a series of models of $N(h)$ above h_{max} and, with the aid of an automatic computer, compute μ_s and ψ_s for each frequency and find by trial and error the $N(h)$ distribution that gives the best agreement with the observations (i).

Acknowledgements—This work was commenced at the Cavendish Laboratory, Cambridge, while we were both on sabbatical leave and we thank Mr. J. A. Ratcliffe for his hospitality and continued interest. We thank also those organizations mentioned in the text who kindly supplied the data we have used. We gratefully acknowledge the assistance of a British Council Scholarship (F. H. H.) and a Nuffield Foundation Fellowship (J. A. T.).

References

Hibberd F. H.	1958	*J. Atmosph. and Terr. Phys.* **12**, 338.
Smart W. M.	1949	*Spherical Astronomy*. Cambridge University Press.
Staff of Mullard Radio Astronomy Observatory	1957	*Nature, Lond.* **180**, 879.
Staff of the Royal Aircraft Establishment	1957	*Nature, Lond.* **180**, 937.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 82 to 85. Pergamon Press Ltd. Printed in Northern Ireland

Sweepers

N. C. Gerson* and W. H. Gossard†

(*Received* 25 *April* 1959)

In a study of atmospherics in the h.f. band, an unusual propagation phenomenon has been observed. This effect seems of sufficient interest to warrant a preliminary report at this time.

The equipment utilized in the experiment consisted of a set of radiofrequency comb filters whose outputs activated a multi-element, moving-strip recorder. By having the threshold of each element slightly above the background noise level, the filter output intensities could modulate the elements. The recorder thus provided a time vs. frequency representation of atmospherics in a given spectral region. Filter bandwidths were all about 8 kc/s, and by employing sixty filters, a total bandwidth of about 500 kc/s was obtained.

With this arrangement, atmospherics occurring in a frequency range $X{\cdot}0$–$X{\cdot}5$ Mc/s could be observed. In this paper the observing frequency, X, will be reported with the understanding that the interval under examination is $X{\cdot}0$–$X{\cdot}5$ Mc/s. The experiment was originally undertaken near Palo Alto, California, so that some check could be made with the whistler measurements underway in that area.

During the observations, it was found that most atmospherics sweep through some finite frequency interval. By far the major proportion of the atmospherics sweep with time from high to low frequencies, although a small fraction run from low to high frequencies. Because of the progression of the radiations across a frequency range, these atmospherics have been termed "sweepers". A summary illustrating the time–frequency variation of the major sweeper types already noted is depicted in Fig. 1.

Sweepers may cover an appreciable portion of the frequency spectrum. In many instances a particular sweeper may pass completely through the 500 kc/s bandwidth under examination with neither its initiation nor termination observed. These radiations and others whose time rate of change equals or exceeds 500 kc/s^2 may be termed "fast sweepers". The rate of change of some fast sweepers exceeded 5×10^3 kc/s^2, the time resolution of the analytical equipment.

Several varieties of sweeper activity appear to be present (see Fig. 1). These may be roughly categorized as normal (frequency decreases with time); reverse (frequency increases with time); fast (frequency change equal to or exceeding 500 kc/s^2); slow (frequency change less than 25 kc/s^2): and irregular (rapid–slow–rapid time rate of frequency change), etc.

In the present experiment, the observing frequency X could be varied from 1–30 Mc/s; however, during the limited observational period sweepers were noted only between 2–30 Mc/s. Results of three morning observations are displayed in Fig. 2. In each instance the results are plotted on a normalized basis, considering a

* Vice Chairman, Arctic Committee, USNC for the IGY.
† U.S. Department of Defence.

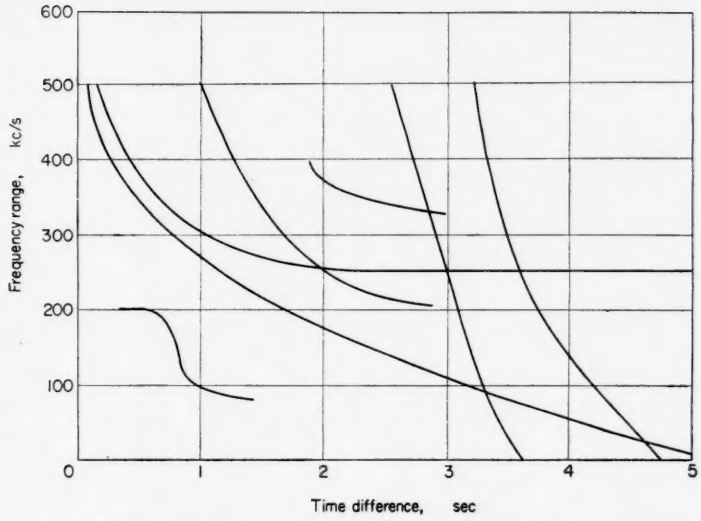


Fig. 1. Principal types of sweeper activity already observed.

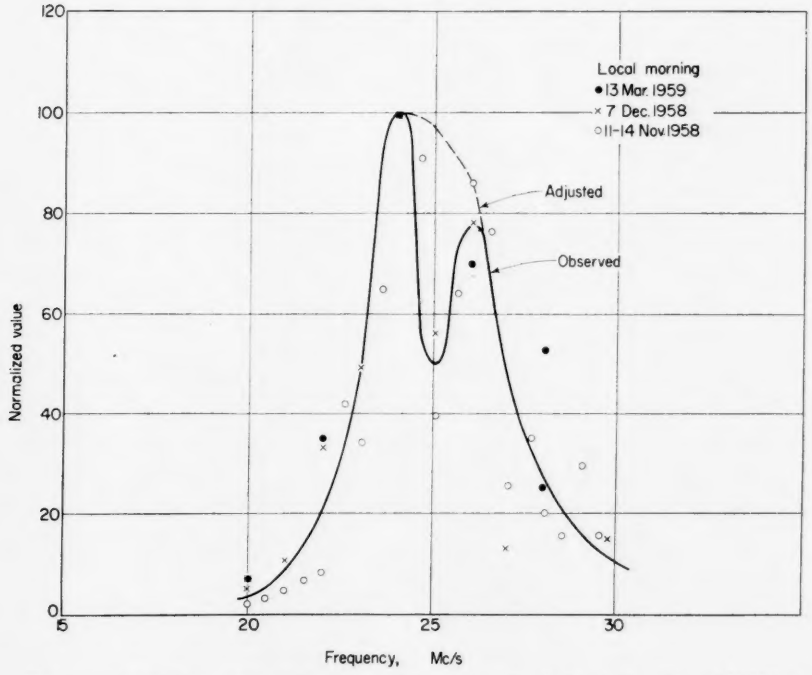


Fig. 2. Distribution of sweepers with frequency normalized at the rate of 2383/hr and 4243/hr for 11–14 November 1958 and 13 March 1959, respectively.

rate of 4243/hr for 13 March 1959, and a rate of 2382/hr for 11–14 November 1958. The maximum occurrence is at 24·0 Mc/s, and the majority, between 23·0–26·5 Mc/s. A tabulation in terms of sweepers/hr for different frequencies in the h.f. band is given in Table 1. It has been noted that the actual number of sweepers/hr at any particular frequency varies from day to day (see Table 1).

A large number of sweepers occur in trains where the individual sweepers recur after a given time interval T. In the observations already made, T varied

Table 1. Rate of sweeper occurrence

Frequency (Mc/s)	Number (hr^{-1})	
	0400–1200 LST 11–14 November 1958	1030–1130 LST 13 March 1959
20·0	18	215
21·0	115	
22·0	208	1504
23·0	810	
24·0	2382	4243
25·0	942	
26·0	2082	2986
27·0	606	
28·0	492	1040–2257
29·0	732	
29·8		643

VOL. 17 1959/6

from 4 sec to a maximum of about 140 sec. In most instances T is not a constant but seems to obey a relationship

$$T = T_0 \pm 0{\cdot}1T_0 \tag{1}$$

However, this relationship is approximate, and further observations are necessary in order to define it more accurately. As an indication of the persistence, one train was observed on fifty-five successive appearances.

For this train, the signal strength of individual sweepers varied from strong to weak (during which one sweeper was missing and one very faintly observable) to strong. The train was observed for 477 sec, and was still being received at the time the test was terminated. The period averaged $T = 8{\cdot}8 \pm 0{\cdot}11$ sec, with a maximum excursion of $8{\cdot}6 < T < 9{\cdot}2$ sec. The sweeper duration was 2·1 sec, the frequency interval 125 kc/s and the initial time rate of change, $\Delta f/\Delta t^2 = 125$ kc/s^2.

Very long sweepers also were observed. In one example, the sweeper duration was about 18 sec, the frequency interval 500 kc/s and the period, T, 140 sec. Another very slow sweeper required 70 sec to progress through 500 kc/s (at a recording frequency of 4 Mc/s).

Attempts to determine a relation between the sweeper duration, d, and the period, T, have thus far not been successful.

A sampling of sweeper activity throughout the 24 hr period indicates that most sweepers observed during local daylight are in the frequency range 20–30 Mc/s. At night the lower frequencies (down to 2 Mc/s) become better observable. During local morning hours sweeper activity is at a maximum between 23–26·5 Mc/s. No

regularity was noted in the initiation or termination frequency, time of onset, duration, or $\Delta f/\Delta t^2$ of particular sweepers during the limited observational programme described here.

With the rather restricted pass band acceptance of most radio receivers it seems tempting to consider the possibility that sweeping atmospherics may be responsible for many of the aural family of onomatopoeia variously known as clicks, tweeks, hisses, swishes, etc., i.e. these phenomena may indicate the audio observance of a limited portion of a sweeper as shaped by the radio receiver. The different sounds then could arise from the acceptance of different portions of the sweeper $\Delta f/\Delta t^2$ curve by the detecting equipment. Near the lower end of its trail, the $\Delta f/\Delta t^2$ becomes smaller and may approach the value of about 1–10 kc/s^2 found for whistlers. It is possible that some of the latter audio atmospherics represent the low frequency portion of sweepers.

To summarize the preliminary results already obtained, the following comments may be made:

(1) Sweepers have been observed within the frequency range 2–30 Mc/s. (However, it is probable that they extend to both higher and lower frequencies. At the lowest frequencies some may be identified with whistlers.)

(2) A bifurcated sweeper occurrence vs. frequency curve was obtained with maxima at 24·0 Mc/s and 26·0 Mc/s. Principal sweeper activity was found between 23·0–26·5 Mc/s.

(3) The occurrence of sweepers in the range 24·0–24·5 Mc/s averaged 1·18/sec between 1030–1130 LST 13 March 1958 and 0·66/sec between 0400–1200 LST (11–14 November 1958). Part of the discrepancy may arise from the greater time interval over which the results were averaged in November 1958. However, the day to day variation in sweeper activity may be considerable.

(4) A diurnal sweeper variation exists in the rate of occurrence and in the distribution by radiofrequency. Sweepers at frequencies below 20 Mc/s are relatively rare, and seem to be best received, during local darkness. Sweepers between 20–30 Mc/s are mainly found during daylight.

(5) The duration of particular sweepers already observed varies from less than 0·1 sec to 18 and even 70 sec.

(6) The frequency coverage of a particular sweeper may vary from 75 kc/s to 1500 kc/s or greater.

(7) Sweepers may appear, on a t–f plot, as t = constant, or as linear, hyperbolic or step functions when f is limited to a range of 500 kc/s (see Fig. 2).

(8) Many sweepers occur in trains where the period (time between successive sweepers of the same train) is approximately constant. Periods already observed vary between 4–140 sec.

(9) The signal intensity of individual sweepers of a train varies from the average. A group of strong intensity sweepers of the same train may be followed by a group of diminishing intensity, thence by a group of increasing signal intensity. The cycle of strong–weak–strong in signal intensity may be repeated.

(10) The source of sweepers may be the sun. Many give the impression of being Type II or Type III solar bursts. If the sun is indeed the source, the reception of low frequency sweepers at night would be indicative of the propagation of rays trapped below the ionosphere.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 86 to 95. Pergamon Press Ltd. Printed in Northern Ireland

The electron density distribution in the *F*-region of the ionosphere

A. J. HIRSH
Cavendish Laboratory, Cambridge

(*Received* 28 *April* 1959)

Abstract—The suggestion of BATES and MASSEY that the *F1*- and *F2*-layers of the ionosphere are formed by a single type of radiation, ionizing an atmosphere in which the electrons are lost by a double process of charge exchange followed by dissociative recombination, is examined in numerical detail. It is shown that under certain circumstances this could lead to the type of $h'(f)$ curve observed in practice. It is also shown that the magnitudes of some of the quantities involved must lie within certain ranges.

1. INTRODUCTION

IT has been suggested that the *F1*- and *F2*-layers of the ionosphere could be produced by one type of incident radiation acting on one atmospheric constituent, if the process of electron loss was a complicated one involving the transfer of charge from an atomic ion to a molecule followed by dissociative recombination of the molecule (BATES and MASSEY, 1946; RATCLIFFE, 1956a). This suggestion is examined here in detail, and we show under what conditions it would lead to the type of ionogram (or $h'(f)$ curve) which exhibits the *F1* critical frequency (f_0F1). The conclusion is reached that although the conditions are somewhat stringent, it is quite possible for them to be satisfied in the actual ionosphere.

VOL.
17
1959/6

2. PRODUCTION AND LOSS OF ELECTRONS

The atmospheric constituent to be ionized is supposed to be an atomic gas (A) with a uniform scale height H, ionized by radiation having the same mass absorption coefficient for all components. The rate (q) of production of electrons in each unit volume when the radiation is from a zenith angle χ is then given by the well-known expression of CHAPMAN (1931):

$$q = q_0 \exp (1 - z - \mathrm{e}^{-z} \sec \chi) \tag{1}$$

where z represents the height, measured in terms of H as a unit, above the level where the peak of q is found when $\chi = 0$.

The electrons are supposed to be lost by processes which involve a molecule (XY) having a scale height equal to H/k so that its number density, $n(XY)$, is given by

$$n(XY) = n_0(XY)\mathrm{e}^{-kz} \tag{2}$$

The electron loss is then supposed to be governed by the reactions

$$A^+ + XY \rightarrow XY^+ + A$$

$$XY^+ + e \rightarrow X + Y$$

The rates of production of A^+ ions and of electrons, respectively, are then given by

$$\frac{d}{dt}n(A^+) = q - \lambda_1 n(A^+)n(XY) \tag{3}$$

$$\frac{d}{dt}n(e) = q - \lambda_2 n(XY^+)n(e) \tag{4}$$

It is also supposed that the atmosphere is electrically neutral, so that

$$n(A^+) + n(XY^+) = n(e) \tag{5}$$

If now it is supposed that the system is in equilibrium, so that

$$\frac{d}{dt}n(A^+) = \frac{d}{dt}n(e) = 0 \tag{6}$$

equations (3) to (6) lead to

$$q = \frac{\lambda_1\lambda_2 n(XY)n^2(e)}{\lambda_1 n(XY) + \lambda_2 n(e)} \tag{7}$$

If $\lambda_1 n(XY) \gg \lambda_2 n(e)$, (7) may be written

$$q = \lambda_2 n^2(e) \tag{8}$$

and if $\lambda_1 n(XY) \ll \lambda_2 n(e)$, it becomes

$$q = \lambda_1 n(XY)n(e) \tag{9}$$

Equation (8) would represent the equilibrium layer of electrons which would be found if the loss process were one of recombination with a constant recombination coefficient α given by

$$\alpha = \lambda_2 \tag{10}$$

Equation (9) would represent the equilibrium layer of electrons which would be found if the loss process were one of attachment with the attachment coefficient β given by

$$\begin{aligned}\beta &= \lambda_1 n(XY) \\ &= \lambda_1 n_0(XY)\mathrm{e}^{-kz}\end{aligned} \tag{11}$$

from (2). Equation (11) implies that the effective attachment coefficient β varies with height. We shall write

$$\lambda_1 n_0(XY) = \beta_0 \tag{12}$$

so that

$$\beta = \beta_0 \mathrm{e}^{-kz} \tag{13}$$

We now solve equation (7) as a quadratic in $n(e)$ to obtain

$$n(e) = \frac{q_0}{2\beta_0}\frac{Q(z,\chi)}{\mathrm{e}^{-kz}}\left[1 + \sqrt{\left\{1 + C\frac{\mathrm{e}^{-2kz}}{Q(z,\chi)}\right\}}\right] \tag{14}$$

where $Q(z, \chi) = \exp(1 - z - \mathrm{e}^{-z}\sec\chi)$ has been put in (1), and where

$$C = \frac{4\beta_0^2}{q_0\alpha} \tag{15}$$

For simplicity, the notation

$$N = N(z) = n(e) \tag{16}$$

will be introduced to represent the variation of electron density with height.

3. Electron Density Distribution

It is now possible to calculate N as a function of z and χ for various values of q_0, α and β_0. It is first necessary to determine the magnitude of k, which is the ratio of the scale height of the atoms A to that of the molecules XY. It has been suggested that A represents atomic oxygen and XY molecular oxygen, in diffusive

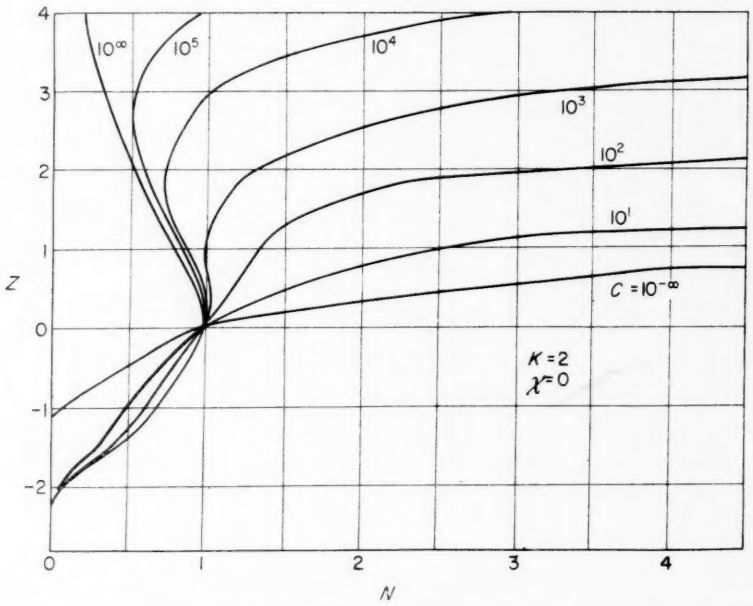


Fig. 1. Electron density distributions for different values of the parameter C, for $k = 2$ and $\chi = 0$.

equilibrium, so that $k = 2$. If, however, as has also been suggested, XY represents molecular nitrogen, and the system is in diffusive equilibrium, then $k = 1{\cdot}75$. If diffusion does not affect the distribution then the two components will have the same scale heights and $k = 1$. If diffusion is only partly effective k may have any value between 1 and either 2 or 1·75. Values of k of 1·2, 1·6 and 2 have therefore been selected to represent some of the different possibilities. In Fig. 1 curves are plotted to show how N varies with height (z) when $k = 2$ and $\chi = 0$. The curves have been normalized and show the ratio of N at height z to N at $z = 0$ plotted against z for different values of the parameter C. The curve for $C = 10^3$ conforms fairly well to what the *F1*-region is supposed to be, and so this case has been investigated more fully in Fig. 2, where the effect of altering χ is shown. For comparison similar curves for $k = 1{\cdot}6$ and $k = 1{\cdot}2$ are shown in Figs. 3 and 4, respectively.

The $h'(f)$ curves corresponding to the $N(z)$ curves of Figs. 2, 3 and 4 were kindly calculated for me by J. E. Titheridge. They are shown in Figs. 5, 6 and 7, respectively, and are seen to be reasonably like those commonly observed.

It is well-known that, if a peak of N is to be produced in the *F2*-region, some

VOL. 17 1959/6

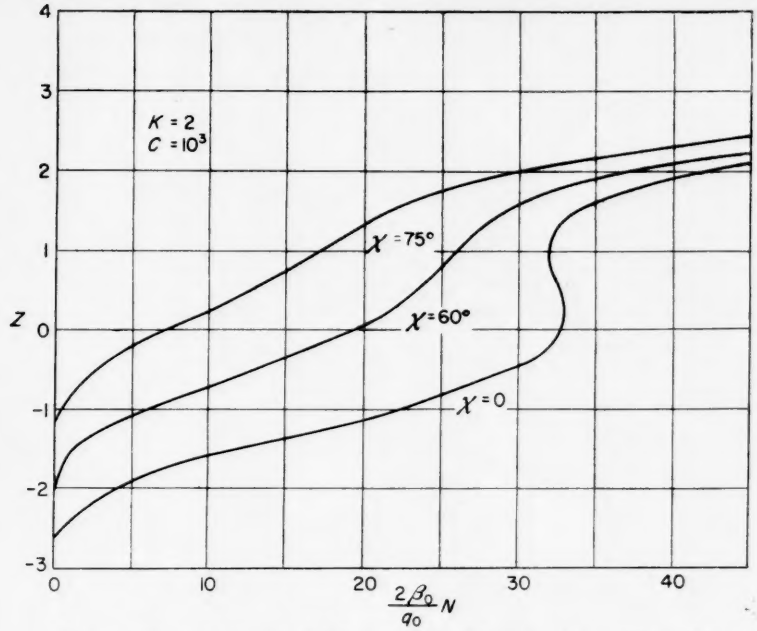


Fig. 2. Electron density distributions for different solar zenith angles χ, for $k = 2$ and $C = 10^3$.

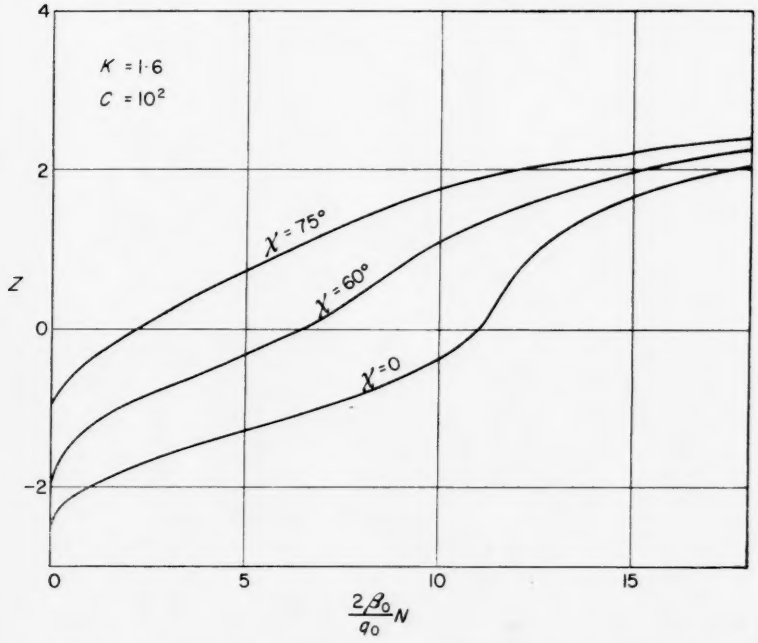


Fig. 3. Electron density distributions for different solar zenith angles χ, for $k = 1{\cdot}6$ and $C = 10^2$.

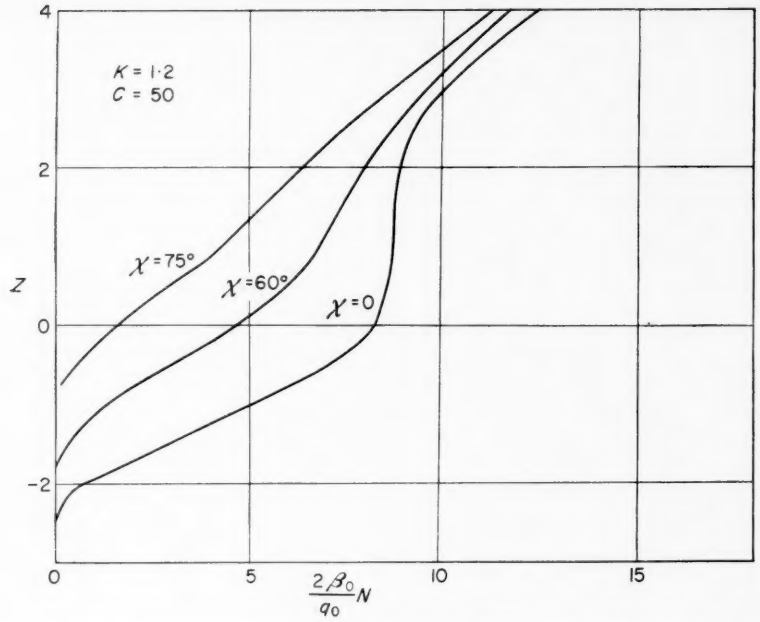


Fig. 4. Electron density distributions for different solar zenith angles χ, for $k = 1{\cdot}2$ and $C = 50$.

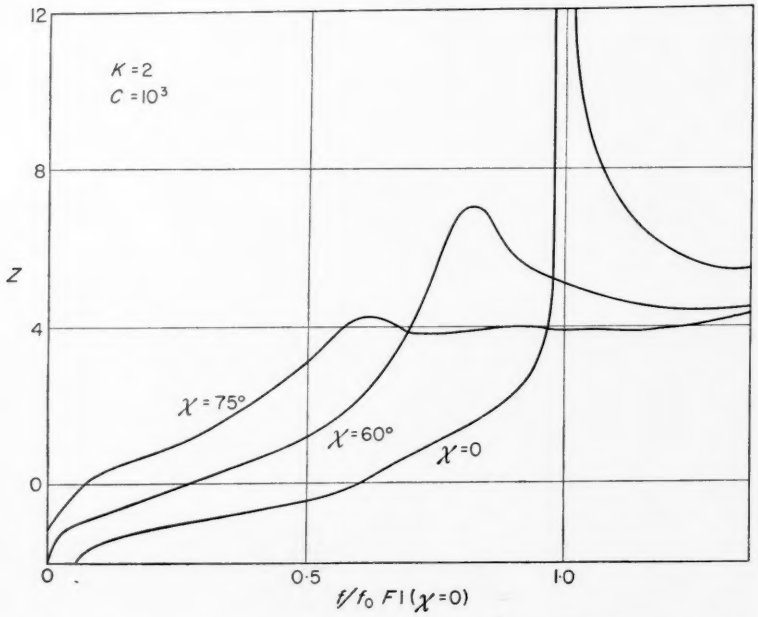


Fig. 5. $h'(f)$ curves corresponding to the electron density distributions in Fig. 2. The frequency scale is normalized to f_0F1 at $\chi = 0$.

VOL.
17
1959/6

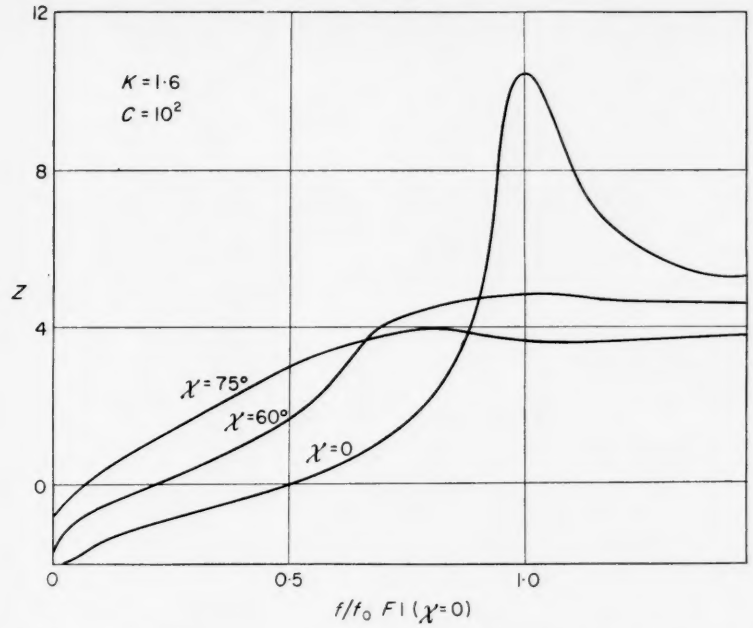


Fig. 6. $h'(f)$ curves corresponding to the electron density distributions in Fig. 3. Scale as in Fig. 5.

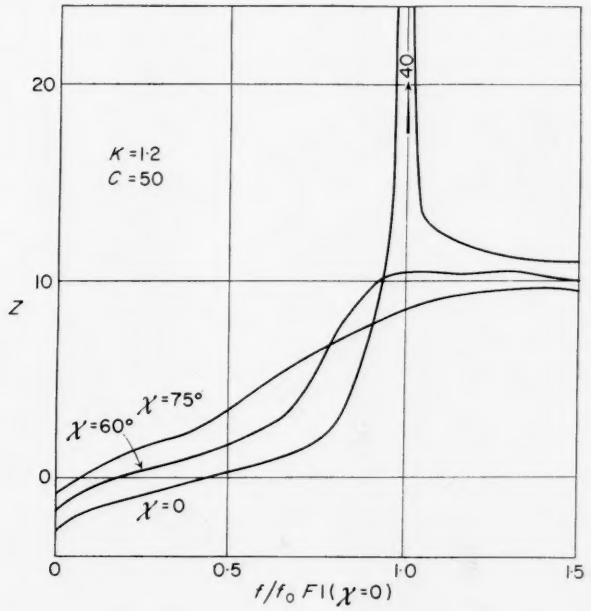


Fig. 7. $h'(f)$ curves corresponding to the electron density distributions in Fig. 4. Scale as in Fig. 5.

other mechanism must be invoked, such as a change in the gradient of β, movements of electromagnetic origin, or diffusion (Yonezawa, 1956). The *F2*-peak is not considered in this paper.

4. Magnitudes of the Parameters

It has been shown that for $k = 2$, when $C = 10^3$, the electron distribution and the $h'(f)$ curves show features characteristic of the *F1*-layer. It must now be asked what other conditions would produce the *F1*-layer and whether these conditions are likely to be found in the ionosphere.

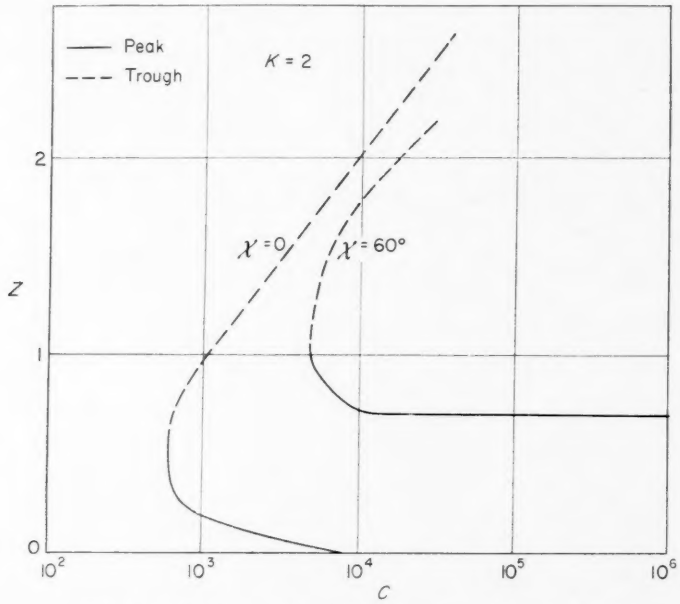


Fig. 8. Effect of C on the heights of peaks and troughs in $N(z)$, for $k = 2$ and for two values of χ.

For this purpose a series of calculations of $N(z)$ was made from equation (14) with different values of C, and Fig. 8 shows the heights of peaks and troughs in N for different values of C, and two values of χ. It is seen that when $\chi = 0$ there is a peak in N only if $C > 600$, and when $\chi = 60°$ only if $C > 5000$. The height, $h_m(N)$, of the peak of N, shown by the continuous lines, is seen to be close to the height, $h_m(q)$, of the peak of q (which is at $z = 0$ when $\chi = 0$, and at $z = 0{\cdot}69$ when $\chi = 60°$), except near the critical values of C. The approximation will therefore be made that $h_m(N) = h_m(q)$, and is at a level where the loss process is effectively one of recombination with the recombination coefficient (α) independent of height. This supposition implies that the peak electron density varies with χ according to

$$N_m = N_0(\cos \chi)^{1/2} \tag{17}$$

where N_0 is the peak electron density when $\chi = 0$, a result known to be approximately, but not exactly, true. Use of (17) allows N_0 to be calculated from values

of N_m determined for known values of χ. Then, with the approximations made here,

$$q_0 = \alpha N_0^2$$

which when combined with (15) gives

$$C = \frac{4\beta_0^2}{N_0^2 \alpha^2} \tag{18}$$

Graphs are now plotted, in the following way, to show the relation between C and β_0. N_0 is calculated by inserting the results of measurement into (17), and

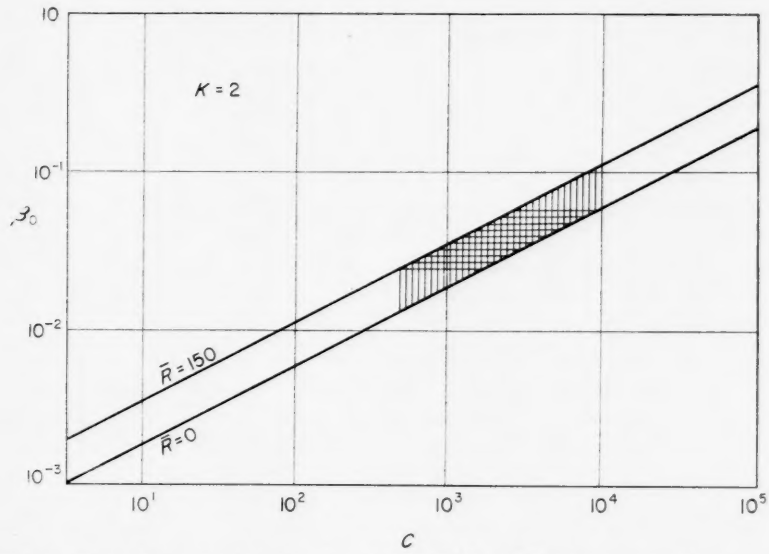


Fig. 9. The relationship between β_0 and C for high and low sunspot numbers, assuming $\alpha = 5 \cdot 10^{-9}$ cm^3-sec^{-1}, showing how limits may be assigned to the value of β_0.

α is assumed to have the magnitude $5 \cdot 10^{-9}$ cm^3-sec^{-1}, which seems, from eclipse measurements (RATCLIFFE, 1956b), to be most likely. The resulting curves are shown in Fig. 9 for sunspot numbers $\bar{R} = 0$ ($N_0 = 2{\cdot}4 \cdot 10^5$ cm^{-3}) and $\bar{R} = 150$ ($N_0 = 4{\cdot}2 \cdot 10^5$ cm^{-3}).

From Fig. 1 it is seen that, if $k = 2$, C must lie in the range $500 < C < 10{,}000$, so that the representative part of the diagram is that shown hatched with vertical lines. It is next supposed that β_0 does not change as the solar cycle advances, so that the representative part is even further restricted to the cross-hatched area on the diagram. This condition restricts the magnitude of β_0 to

$$2{\cdot}5 \cdot 10^{-2} \text{ sec}^{-1} < \beta_0 < 5{\cdot}6 \cdot 10^{-2} \text{ sec}^{-1} \tag{19}$$

Corresponding limits for $k = 1{\cdot}6$ are

$$1{\cdot}0 \cdot 10^{-2} \text{ sec}^{-1} < \beta_0 < 1{\cdot}8 \cdot 10^{-2} \text{ sec}^{-1} \tag{20}$$

and for $k = 1{\cdot}2$ they are

$$3{\cdot}5 \cdot 10^{-3} \text{ sec}^{-1} < \beta_0 < 5{\cdot}6 \cdot 10^{-3} \text{ sec}^{-1} \tag{21}$$

Thus it is seen that a knowledge of the value of β_0 would enable the value of k to be fixed and would also fix the range over which C varies with q_0.

In the paper by Ratcliffe *et al.* (1956) some magnitudes of β_0 were suggested, and, although their conclusions depend on some hypotheses which are not necessarily correct, it is interesting to compare the present result with theirs. They suggested that

$$\beta = 10^{-4} \exp \frac{300 - h}{45} \text{ sec}^{-1}$$

and that $h_0(q) = 180$ km. This suggestion would lead to $\beta_0 = 1{\cdot}5 \,.\, 10^{-3}$ sec^{-1}.

5. Non-Chapman Behaviour of f_0F1

In the discussion of Fig. 9 it has been supposed that f_0F1 behaves as though it corresponded to the peak of an α-Chapman layer with

$$f_0F1 \propto (\cos \chi)^n \tag{22}$$

and $n = 0{\cdot}25$. It is therefore of interest to enquire what value of n would be required in equation (22) if it were to represent the results of the hypothetical layer considered in this paper. From the series of $h'(f)$ curves computed by Titheridge from theoretical $N(z)$ curves calculated for different values of χ, with $C = 10^3$ and $k = 2$, it is found empirically that $n = 0{\cdot}2$. It is interesting to note that the experimentally determined values of n lie between 0·28 and 0·13, with a mean value of 0·20 (Allen, 1948; my n is the reciprocal of his).

The non-Chapman behaviour of f_0F1 makes only a small change in the results of Fig. 9 for the magnitude deduced for β_0.

6. Conclusions

The theory of a double loss process as described in Section 2 leads to the sort of $h'(f)$ record found experimentally. The value deduced for n in equation (22) is roughly the same as the value found experimentally. The theory shows how the loss process changes from being recombination-like to being attachment-like as z increases.

The theory in Section 4 shows how the values of some of the parameters involved may be evaluated within fairly narrow ranges. Detailed examination of the non-Chapman behaviour of f_0F1 may help to restrict these values still further.

Acknowledgement—I am indebted to Mr. J. A. Ratcliffe for suggesting this problem to me and for his help and advice in the preparation of this paper. I wish to express my thanks for the award of a McGill Delta Upsilon Memorial Scholarship which has enabled me to carry out this work.

Note added in proof

While this paper was going to press my attention was drawn to a similar paper just published by Yonezawa *et al.* (1959).

References

Allen C. W.	1948	*J. Terr. Mag. Atmos. Elect.* **53,** 433.
Bates D. R. and Massey H. S. W.	1946	*Proc. Roy. Soc.* A **187,** 261.
Chapman S.	1931	*Proc. Phys. Soc.* **43,** 26.

VOL. 17 1959/6

Ratcliffe J. A.	1956a	*J. Atmosph. Terr. Phys.* **8,** 260.
Ratcliffe J. A.	1956b	*J. Atmosph. Terr. Phys.* Special Supplement on Solar Eclipses and the Ionosphere, 1.
Ratcliffe J. A., Schmerling E. R., Setty G. S. K. and Thomas J. O.	1956	*Phil. Trans.* A **248,** 621.
Yonezawa T.	1956	*J. Radio Res. Lab.* **3,** 1.
Yonezawa T., Takahashi H. and Arima Y.	1959	*J. Radio Res. Lab.* **6,** 21.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 96 to 109. Pergamon Press Ltd. Printed in Northern Ireland

The calculation of real and virtual heights of reflection in the ionosphere

J. E. Titheridge
Cavendish Laboratory, Cambridge

(*Received* 29 *April* 1959)

Abstract—A rapid and accurate manual method is described by means of which the heights h corresponding to a given series of electron densities N can be calculated from an ionogram which shows the $h'(f)$ curve for the ordinary or extraordinary wave. The method makes allowance for the presence of the earth's magnetic field. The virtual height is read only once at about twenty frequencies, and the calculation of a complete $N(h)$ curve requires less than 15 min. A slightly modified method is described for use when very accurate results are required, as, for example, in a study of the fine structure of the E-layer.

The method makes use of a series of coefficients which may be quickly calculated once and for all, for a given place, with the aid of a desk calculating machine. The law assumed for the shape of the segments used in the analysis eliminates the necessity for the calculation and subsequent integration of the group refractive index in deriving these coefficients. The coefficients for the extraordinary wave are readily obtained by applying a correction to the "longitudinal" expression for the group refractive index.

It is shown how the same coefficients can be used in the inverse process of deriving an $h'(f)$ curve from a known $N(h)$ curve.

1. Introduction

Experimental observations of the variation of the virtual height of reflection (h') of a radio wave transmitted vertically, as a function of the wavefrequency (f), provide an extremely important source of information about the distribution of electron density (N) with height (h) in the ionosphere. Only in the last few years, however, have suitable methods (including the effects of the earth's magnetic field) been developed for analysing these $h'(f)$ records to give this information. A survey of both digital computer and manual methods of doing this calculation has been given recently (Thomas, 1959).

In this paper a method is presented which has some advantages over those at present in use, for the rapid and accurate analysis of ionospheric records without the use of mechanical aids. It differs from existing methods by working in terms of the *changes* in height from one frequency to the next, and of the *changes* in the group retardation due to a given segment of ionization as the wave frequency is increased. It is equally suitable for the analysis of ordinary and extraordinary wave traces.

The time required to analyse a given $h'(f)$ record is considerably less than that required by the manual methods now in use, since the $h'(f)$ curve need only be read once at about twenty given frequencies and the complete $N(h)$ curve can then be calculated in under 15 min (Section 2.2). The present method is particularly suitable for the analysis of accurate $h'(f)$ records used in the study of the fine structure of the ionosphere. In a simplified form the method may be used for rapid routine analysis without the use of any computing aids (Section 2.3).

It is also possible to calculate the virtual heights corresponding to a series of known real heights, using the same coefficients as in the analysis of a virtual height record (Section 2.4). This is particularly important in a procedure described in a

companion paper (TITHERIDGE, 1959) which uses both the ordinary and extraordinary ray $h'(f)$ data to determine the amount of ionization below the lowest recorded virtual height, the size of any "valley" or minimum of ionization between two layers, and the errors in the $N(h)$ curves (as normally calculated) due to the neglect of these two regions. The method of analysis described in the present paper also considerably simplifies these calculations by permitting the calculation of a complete $N(h)$ curve when the real heights are known (or are postulated) up to some particular plasma frequency, and the virtual heights are known above this frequency (Section 2.2).

2. The Method of Analysis

2.1. *Principle*

The virtual height of reflection of a vertically incident radio wave is given by

$$h' = \int_0^h \mu' \, dh \tag{1}$$

where the group refractive index μ' is a function of the electron density N, the

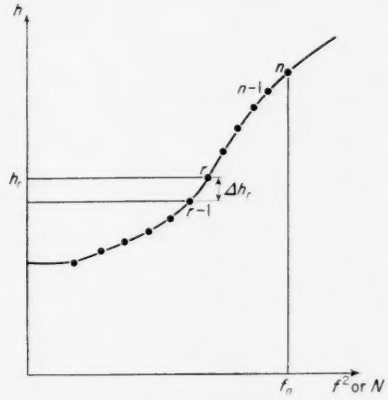


Fig. 1. The notation used in the sub-division of an $N(h)$ curve.

wavefrequency f, and the strength and direction of the earth's magnetic field.

The $N(h)$ curve may be thought of as being divided up into a series of heights h_r (Fig. 1) corresponding to the virtual heights of reflection h_r' at a given series of frequencies f_r. If $\bar{\mu}_{rn}'$ is the mean value of μ' over the lamination h_{r-1} to h_r, for a wave of frequency f_n, then

$$\bar{\mu}_{rn}' = \frac{1}{h_r - h_{r-1}} \int_{h_{r-1}}^{h_r} \mu' \, dh \tag{2}$$

In order to evaluate $\bar{\mu}'$ for any lamination some assumption will have to be made about the shape of the $N(h)$ curve. It is normally assumed that h varies linearly with N or f within a particular lamination; it will, however, be shown in Section 3 that an alternative variation can be found which considerably reduces the work involved in the calculation of $\bar{\mu}'$.

Since the virtual height at any frequency is equal to the real height at that frequency plus the sum of the group retardations produced by each of the underlying laminations, we have

$$\begin{aligned} h_0' &= h_0 \\ h_1' &= h_1 + (\bar{\mu}_{11}' - 1)(h_1 - h_0) \\ h_2' &= h_2 + (\bar{\mu}_{22}' - 1)(h_2 - h_1) + (\bar{\mu}_{12}' - 1)(h_1 - h_0) \\ &\quad \text{etc.} \end{aligned} \tag{3}$$

When the virtual heights h_0', h_1' ... are known, the successive real heights h_0, h_1, ... can be found by solving these equations in turn. This is the method employed in the machine calculation of THOMAS *et al.* (1958) and in the manual methods of JACKSON (1956) [using the equations in the form $h_n' = h_0 + \sum_{r=1}^{r} \bar{\mu}'_{r,n}(h_r - h_{r-1})$].

Alternatively the set of linear equations can be inverted to give the real heights directly in terms of the known virtual heights (BUDDEN, 1954).

If, however, we write $h_1 - h_0 = \Delta h_1$, $h_2 - h_1 = \Delta h_2$, etc., equations (3) give, by subtraction,

$$\begin{aligned} h_0' &= h_0 \\ h_1' - h_0' &= \Delta h_1' = \Delta h_1 + (\bar{\mu}_{11}' - 1)\,\Delta h_1 = \bar{\mu}_{11}'\,\Delta h_1 \\ h_2' - h_1' &= \Delta h_2' = \{h_2 + (\bar{\mu}_{22}' - 1)\,\Delta h_2 + (\bar{\mu}_{12}' - 1)\,\Delta h_1\} - \\ &\qquad\qquad - \{h_1 + (\bar{\mu}_{11}' - 1)\,\Delta h_1\} \\ &= \bar{\mu}_{22}'\,\Delta h_2 - (\bar{\mu}_{11}' - \bar{\mu}_{12}')\,\Delta h_1 \\ &\quad \text{etc.} \end{aligned} \tag{4}$$

Writing

$$\begin{aligned} \beta_{r,n} &= \bar{\mu}'_{r,n-1} - \bar{\mu}'_{r,n}\ (r < n) \\ \beta_{n,n} &= \bar{\mu}'_{n,n} \end{aligned} \tag{5}$$

the set of equations (4) may be solved successively to give

$$\begin{aligned} h_0 &= h_0' \\ h_1 - h_0 = \Delta h_1 &= \frac{1}{\beta_{11}}\,\Delta h_1' \\ h_2 - h_1 = \Delta h_2 &= \frac{1}{\beta_{22}}\,\{\Delta h_2' + \beta_{12}\,\Delta h_1\} \\ h_3 - h_2 = \Delta h_3 &= \frac{1}{\beta_{33}}\,\{\Delta h_3' + \beta_{13}\,\Delta h_1 + \beta_{23}\,\Delta h_2\} \end{aligned}$$

and in general

$$\Delta h_n = \frac{1}{\beta_{nn}}\,\{\Delta h_n' + \sum_{r=1}^{n-1} \beta_{rn}\,\Delta h_r\} \tag{6}$$

The calculation of the values of Δh_n from this relation, using the known values of $\Delta h_n'$, is essentially similar to the calculation of successive values of h_n, using the

VOL. 17 1959/6

known values of h_n', from equations (3). In practice, however, the present method offers two advantages:

(a) Δh_n is normally only a few kilometres and is given by equation (6) as the *sum* of a number of calculated terms. Since the accuracy required is only, say, 0·2 km, no part of the calculation requires an accuracy of greater than 5 or 10 per cent.

(b) The coefficients β_{rn} as defined by equation (5) represent the *decrease* in the retardations in a given block as the wavefrequency is increased from f_{n-1} to f_n. They are thus negligibly small when $r \ll n$ (Table 1) even although the coefficients $(\bar{\mu}' - 1)$ of equation (3) may still be appreciable. So even at $n = 20$ the summation in equation (6) need include only four or five terms to give a value of Δh accurate to within 1 per cent.

An $N(h)$ curve can thus be accurately obtained from equation (6) by a very simple calculation requiring no mechanical aids (except, perhaps, a slide-rule) in which each value of Δh is obtained from the sum of about four numbers each of one or two digits. Further, the calculations can be so arranged (Section 2.2 and Table 2) that these four numbers are, in effect, obtained from a single setting of a slide-rule.

2.2. *Practical calculations*

The coefficients β shown in Table 1 were calculated by the method of Section 3.1 for E-region analysis in south-east England (dip angle = 67°, gyrofrequency = 1·2 Mc/s). The frequencies used are given by

$$f_n = \frac{1{\cdot}5}{\log 1{\cdot}5} \log \left(1 + \frac{n}{2}\right) \text{ Mc/s}$$

and were chosen to provide nearly equal real height intervals Δh over the complete $N(h)$ curve. These coefficients may be used directly in equation (6) to calculate the real heights h_n at the above frequencies. Thus the real height h_6, at a frequency of 2·77 Mc/s, is obtained from

$$10{\cdot}85(h_6 - h_5) = 10{\cdot}85\ \Delta h_6 = \Delta h_6' + 6{\cdot}98\ \Delta h_5 + 0{\cdot}25\ \Delta h_4 + \\ + 0{\cdot}06\ \Delta h_3 + 0{\cdot}02\ \Delta h_2.$$

It will be seen that the last terms contribute only a very small amount to the total. Thus if Δh_2, Δh_3 and Δh_6 are each about 1 km, neglecting the last two terms in the above relation gives a value of Δh_6 which is too small by only 0·008 km, or less than 1 per cent. The coefficients above the solid line in Table 1 have values which are less than 1 per cent of the diagonal value β_{nn}, and the sum of all the coefficients in any column above this line is about 1·5 per cent of β_{nn}. Since the steps Δh are all approximately equal with the frequency scale used, the upper coefficients in Table 1 may be completely neglected (i.e. put equal to zero) without introducing errors of more than 2 or 3 per cent in the calculated values of Δh.

This error is considerably reduced, and given a random sign, if the lower coefficients of Table 1 are increased to compensate for the neglected upper coefficients, so that the value of $\sum_{r=1}^{n-1} \beta_{rn}$ is unchanged. The value of Δh_n obtained

Table 1. The coefficients β for the case of twenty-five laminations at frequencies proportional to log $(1 + n/2)$.

$n =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
$f =$	1·50	1·96	2·26	2·47	2·64	2·77	2·89	2·99	3·08	3·15	3·22	3·29	3·35	3·40	3·45	3·50	3·54	3·57	3·61	3·65	3·68	3·71	3·74	3·77
$r = 2$		3·24	1·98	0·10	0·04	0·02	0·02	0·02	0·01	0·01	0·01	0·01	0·01	0	0	0	0	0	0	0	0	0	0	0
3			5·15	3·68	0·17	0·06	0·03	0·02	0·02	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0	0	0
4				7·07	5·36	0·25	0·10	0·05	0·04	0·03	0·02	0·02	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01	0·01
5					8·96	6·98	0·37	0·13	0·06	0·05	0·04	0·03	0·02	0·02	0·02	0·02	0·02	0·02	0·01	0·01	0·01	0·01	0·01	0·01
6						10·85	8·57	0·52	0·17	0·06	0·05	0·04	0·03	0·03	0·03	0·02	0·02	0·02	0·02	0·02	0·01	0·01	0·01	0·01
7							12·57	10·13	0·71	0·23	0·08	0·05	0·04	0·04	0·04	0·03	0·03	0·02	0·02	0·02	0·02	0·02	0·01	0·01
8								14·55	11·6	0·90	0·28	0·12	0·07	0·06	0·05	0·04	0·03	0·02	0·02	0·02	0·02	0·02	0·02	0·02
9									16·35	12·9	1·12	0·34	0·15	0·11	0·08	0·05	0·04	0·03	0·03	0·02	0·02	0·02	0·02	0·02
10										18·05	14·2	1·35	0·45	0·18	0·14	0·09	0·05	0·04	0·03	0·03	0·03	0·02	0·02	0·02
11											19·8	15·5	1·60	0·58	0·21	0·16	0·10	0·07	0·06	0·05	0·04	0·03	0·02	0·02
12												21·55	16·7	1·86	0·68	0·26	0·17	0·10	0·07	0·07	0·06	0·05	0·04	0·03
13													23·2	17·8	2·12	0·77	0·33	0·18	0·10	0·08	0·07	0·06	0·05	0·04
14														24·7	18·8	2·40	0·87	0·39	0·20	0·12	0·08	0·07	0·06	0·05
15															26·2	19·7	2·70	0·94	0·45	0·24	0·14	0·10	0·06	0·06
16																27·8	20·6	3·00	1·01	0·52	0·28	0·16	0·10	0·07
17																	29·3	21·5	3·30	1·11	0·59	0·32	0·20	0·10
18																		30·7	22·3	3·58	1·22	0·64	0·34	0·22
19																			32·0	23·1	3·81	1·40	0·70	0·36
20																				33·3	23·8	4·00	1·56	0·80
21																					34·6	24·6	4·17	1·69
22																						35·9	25·3	4·27
23																							37·2	26·1
24																								38·6

VOL. 17
1959/60

from equation (6) is then almost unchanged when the steps Δh are approximately equal. Similarly the remaining coefficients can be rounded off to values suitable for rapid calculation provided that their sum is unchanged. If the diagonal coefficient β_{nn} is rounded off the sum of the other coefficients must be changed by an equal amount.

The modified coefficients obtained in this way are shown in Table 2. The "diagonal" coefficients β_{nn} are given here in the bottom line of the table, with the (modified) values of $\beta_{n-1,n}$; $\beta_{n-2,n}$, etc., in the column immediately above. With a table laid out in this manner, the analysis can be performed in a particularly rapid manner. The calculations are carried out on a sheet of semi-transparent paper placed over the table, calculated terms being written on top of the factors used. The observed values of h' are inserted in the table in the position indicated (underneath the corresponding frequency) and the differences $\Delta h'$ are evaluated. The first "total" is simply the first difference $\Delta h_1'$, and is divided by 1·2 to give the corresponding real height step Δh_1. The values of 0·18 Δh_1 and 0·01 Δh_1 are then written over the factors 0·18 and 0·01. The second total ($\Delta h_2' + 0{\cdot}18\ \Delta h_1$) is then obtained, giving (on division by 3·24) the second step Δh_2 and its multiples 2 Δh_2 and 0·1 Δh_2 contributing to the third and fourth totals. Successive steps Δh are thus rapidly developed, being finally combined to give the corresponding values of h.

To calculate about twenty points each addition consists of only three or four terms (of one or two digits each) and the same slide-rule setting used to divide this sum by the appropriate factor to get Δh is used to read off the two or three multiples of Δh required for succeeding additions (this avoids the carrying forward of rounding-off errors). This is the only work required for each step, and an analysis may be completed in about 10 min. The steps of from $\frac{1}{2}$ to 1 km are obtained to within 0·01 km for a normal E-layer.

In Table 2 the first virtual height (at $n = 0$) corresponds to an electron density of about 3000 electrons/cm^3 and is found from low frequency experiments or adjusted empirically to give the correct amount of low-level ionization. From this point the true height is supposed to vary linearly with $N^{-\frac{1}{2}}$ (Section 3) up to the second point at $f = 1{\cdot}5$ Mc/s. This part of the curve is shown dashed in Fig. 2. The figures in small type at the foot of Table 2 are used to provide a check on the accuracy of the calculations, and are discussed in Section 2.4.

Table 2 may also be used directly to calculate a complete $N(h)$ curve when the $h'(f)$ curve is only known down to some middle frequency, and some real height model is assumed below this frequency. The known steps Δh are simply inserted in the table and the calculation carried on in the normal way.

Tables similar to Table 2 are available from the author for both the ordinary and extraordinary rays, with linear and logarithmic frequency scales and maximum frequencies of 2 to 7 Mc/s, calculated for a dip angle of 67° and for conditions at the magnetic equator.

2.3. *Routine calculations of real heights*

Since $\bar{\mu}'$ does not vary rapidly with the ratio Y of the gyrofrequency to the wavefrequency, only the ratios of the frequencies at which the $h'(f)$ curve is read

Table 2. The modified coefficients arranged in a form particularly suitable for rapid real height calculations using a slide-rule.
The working is carried out on a sheet of semi-transparent paper placed over the table.

h																								
n																								
f																								
h'																								
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
	—	1·50	1·96	2·26	2·47	2·64	2·77	2·89	2·99	3·08	3·15	3·22	3·29	3·35	3·40	3·45	3·50	3·54	3·57	3·61	3·65	3·68	3·71	3·74	3·77
																		0·5	0·5	0·5	0·7	1	1	1·5	1·3
											0·3	0·5	0·5	0·4	1	1	1·25	1	1	1	1	1·6	1·6	2	2
											1	1	1·5	2	2	2·5	2·5	3	3	4	4	4	4	4	4·5
											13	14·5	15	16·5	18	18·5	20	21	22	22	23	24	25	25	26·5
											18	20	21	23	25	26	28	30	31	32	33	35	36	37	39

need be exactly specified. Thus if the curve is available with a logarithmic frequency scale, and a transparent scale with vertical lines corresponding to the desired reading points is used, this scale may be moved horizontally into any desired position. For maximum accuracy the reading points should be more closely spaced at higher frequencies where the slope of the $h'(f)$ curve is greatest. The number of lines crossing the curve, i.e. the number of points to be read, can then be varied by shifting the scale. The position can also be adjusted to record cusps on the $h'(f)$ record which might not be adequately recorded when reading the curve at a fixed series of points. This in turn means that, for a given accuracy, fewer points are required. This procedure is similar to that adopted by KELSO (1954) using a 5-point integration relation (later modified by SHINN (unpublished) and SCHMERLING (1958) to include the effects of the magnetic field). However, since in the present method the $h'(f)$ curve is only read once, we do not get the spurious irregularities introduced by the 5-point methods whenever large cusps are present on the $h'(f)$ record.

A given table is accurate over only a small range of frequencies, or positions of the reading grid, due to the change in Y. The error increases almost linearly with the change in frequency, and three tables are required to cover a two-to-one frequency range with a maximum error of 2 per cent in Δh. A much greater range can of course be covered if the number of points read is varied, or if the permissible error is greater.

The full accuracy given by Table 2 is only justified when the virtual heights can be measured to within 0·5 km, and the first few to within 0·2 km. For normal calculations the table of coefficients may be further simplified (by the methods discussed in the previous section) to the form shown in Table 3, which neglects all the coefficients β_{rn} which have a value of less than 0·04 β_{nn}. For $r < n$ the values actually tabulated are β_{rn}/β_{rr} so that the successive totals $\beta_{nn}\,\Delta h_n$ are found by adding the products of these tabulated factors with the previous totals (equation 6). In this form the factors are more suitable for mental calculation.

In use the differences $\Delta h'$ are first entered in the table, and the successive totals T_1, T_2, T_3, etc., obtained from the relations

$$\begin{aligned} 1{\cdot}2(h_1 - h_0) &= T_1 = \Delta h_1' \\ 3(h_2 - h_1) &= T_2 = \Delta h_2' + \tfrac{1}{7}T_1 \\ 5(h_3 - h_2) &= T_3 = \Delta h_3' + 0{\cdot}6T_2, \qquad \text{etc.} \end{aligned}$$

Finally these totals are divided by the factors 1·2, 3, 5, . . . to give the required values of Δh_1, Δh_2, Δh_3, ... The complete process takes perhaps 5 min.

The coefficients of Table 3 look very approximate. The round numbers were, however, obtained by carefully modifying the exact coefficients in the manner discussed in Section 2.2, and are in reality quite accurate. For calculations on a single layer, when the values of Δh are not large and variable, the true heights obtained are accurate to within about 0·2 km. As an example, an irregular $h'(f)$ trace for the E-layer was analysed using Table 2, the real heights being calculated to two decimal places. The virtual height curve used (recorded at Cambridge at

Table 3. A simplified form of Table 2, for routine *E*-layer calculations without the use of a slide-rule.

h																									
n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
h'																									
$\Delta h'$									0·1	0·1	0·1	0·1	0·1	0·1	0·1	0·3	0·2	0·2	0·1	0·2	0·2	0·2	0·2	0·2	0·2
			1/7	0·6	0·8	0·9	0·7	0·7	1	0·7	1	0·7	0·7	0·7	0·7	1	0·7	0·7	0·7	0·7	0·7	1	0·7	0·7	0·7
Total																									
Δh	1·2	3	5	7	10	10	10	15	15	15	20	20	20	20	20	30	30	30	30	30	30	40	40	40	40

VOL. 17 1959/6

0520 on 18 June 1956) is shown by the solid line in Fig. 2, while the dots show the calculated real heights. The real heights were also calculated (from the same values of h') using Table 3. The difference between the values of h_n calculated by the two methods had a maximum value of 0·18 km, and a mean value of 0·07 km. With a normal $h'(f)$ curve the error is even less than this.

The form assumed for the $N(h)$ curve at low values of N is shown by the broken

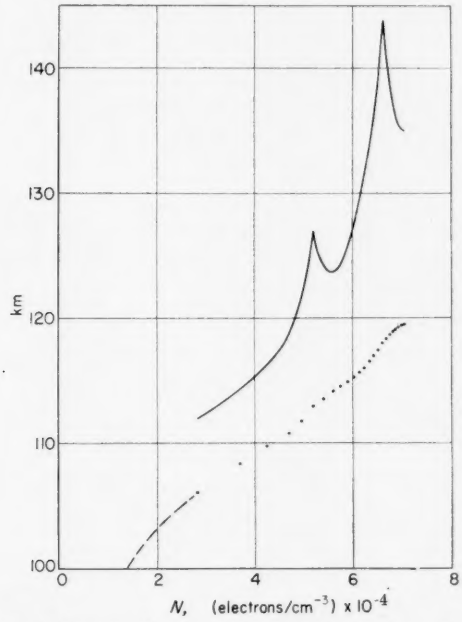


Fig. 2. An irregular $h'(f)$ curve for the E-layer (upper line) as observed at Cambridge at 0520 on 18 June 1956. The dots show the calculated "real" heights, and the dashed line shows the allowance for the low-level ionization.

line in Fig. 2. It is a linear variation of h with $N^{-\frac{1}{2}}$ down to a height of 76 km (the observed height of reflection at a frequency of 15 kc/s) at $N = 3000$ electron/cm^3.

2.4. *Virtual-height calculations*

Table 2 can also be used to calculate the virtual heights corresponding to a known real height curve. The values of Δh are now known and all the terms $\Delta h_r \beta_{rn}$ can be filled in. The values of $\Delta h'$ are then given by the amount required to give the correct total $\Delta h_n \beta_{nn}$.

If h' is required at only a few frequencies, it can be obtained directly from the relation

$$h_n' = h_n + \sum_{r=1}^{n-1} (\bar{\mu}_{rn}' - 1)\, \Delta h_r \qquad (7)$$

This relation may also be used to check that no errors have been made in the real height calculation carried out in the way already described. For suppose that a calculating error causes a value of Δh_r to be 0·1 km too large. Then the values of

Δh_{r-1}, Δh_{r+2}, . . . calculated from equation (6) are also increased, but only by progressively smaller amounts so that the calculated real height h_n at some higher frequency will be perhaps 0·3 km too large. Thus every term in equation (7) is increased and the calculated value of h_n' will be about 2 km greater than the true value. A comparison of the virtual height calculated at a single frequency with the value read off the $h'(f)$ record thus provides a very sensitive check on the accuracy with which the preceding steps in the calculation were carried out. The factors $(\bar{\mu}_{rn}' - 1)$ for $n = 15$ are given at the bottom of Table 2. The terms $(\bar{\mu}_{rn}' - 1)\,\Delta h_r$ are obtained in the course of the analysis when the terms $\beta_{rn}\,\Delta h_r$ are being calculated, and are later summed to provide an over-all check.

When the extraordinary ray virtual heights corresponding to a known ordinary ray $h'(f)$ curve are to be calculated, it is convenient to place the coefficients $(\bar{\mu}' - 1)$ for this ray (calculated at frequencies which reach the same real height as the ordinary ray frequencies used) as a continuation of the table of coefficients β used in the ordinary ray analysis. The $N(h)$ curve and the corresponding extraordinary ray $h'(f)$ curve are then calculated in one operation. The virtual heights are obtained from equation (7) and are largely independent of any rounding off in the calculated true heights. This procedure is the basis of the calculations described in the companion paper (TITHERIDGE, 1959).

3. CALCULATION OF THE COEFFICIENTS β

The coefficients β are evaluated for a given place once and for all as a preliminary calculation in the way described below. In previous work, h has been taken to vary linearly with the plasma frequency $f_n[= \sqrt{(Ne^2/\pi m)}]$ (BUDDEN, 1954; KING, 1954; THOMAS *et al.*, 1958) or linearly with N (JACKSON, 1956) over the elementary sections Δh. In the latter case (2) becomes

$$\bar{\mu}_{rn}' = \frac{1}{\Delta N}\int_{N_{r-1}}^{N_r} \mu'(N)\,dN$$

μ' may be obtained with the aid of an electronic computor, or by substituting calculated values of μ (equation 10) in the relation

$$\bar{\mu}' = \mu + f\frac{\partial \mu}{\partial f} \tag{8}$$

the differential being determined graphically. The resultant $\mu'(N)$ curve is then integrated numerically up to various values of N, and the calculations repeated for different frequencies.

This work can be considerably reduced if h is assumed to vary linearly with $X^{-\frac{1}{2}} = w$, say, (where $X = Ne^2/\pi m f^2$) over the elementary segments into which the $N(h)$ curve is divided. At a given location μ is a function of w and Y only (equation (10) and equation (2) gives

$$\bar{\mu}' = \frac{1}{\Delta w}\int_{w}^{w+\Delta w}\left(\mu + f\frac{\partial \mu}{\partial f}\right)dw$$

$$= \frac{1}{\Delta w}\int_{w}^{w+\Delta w}\left(\mu + w\frac{\partial \mu}{\partial w} - Y\frac{\partial \mu}{\partial Y}\right)dw$$

VOL. 17 1959/6

or
$$\bar{\mu}' = \frac{\Delta(\mu w)}{\Delta w} - \overline{Y \frac{\partial \mu}{\partial Y}} \tag{9}$$

It is now only necessary to calculate μ at the end points of each interval. The Appleton–Hartree relation is used for calculating μ and may be written in the form
$$\frac{X}{1 - \mu^2} = 1 + Y \cos\theta - S \pm \sqrt{(1 + S^2)} \tag{10}$$

where
$$S = \frac{Y^2 \sin^2\theta}{z(1 - X)\cos\theta}$$

and θ is the acute angle between the wave-normal and the direction of the magnetic field. The upper and lower alternative signs refer to the ordinary and extraordinary rays, respectively.

3.1. *Ordinary wave*

For the ordinary wave, the second term in (9) is less than 5 per cent of $\bar{\mu}'$ and is fairly constant. It is readily obtained from the calculated values of μ and plotted as a function of w at several frequencies.

If, however, a mean error of about 1 per cent can be tolerated, it is only necessary to use the approximate relation
$$\bar{\mu}' = 0{\cdot}98 \frac{\Delta(\mu w)}{\Delta w}$$
instead of the expression (13).

The difference between the segments assumed above and those in which h varies linearly with $\sqrt{N}$ or N is negligible when the step ΔN is less than $N/5$.

3.2. *Extraordinary wave*

For the extraordinary ray, $Y\,\partial\mu/\partial Y$ is not well-behaved and becomes infinite at reflection. It is then preferable to plot $2\sqrt{(w - w_0)}.\ Y\partial\mu/\partial\ Y$ against $\sqrt{(w - w_0)}$, where w_0 is the value of w at reflection. $\overline{Y\ \partial\mu/\partial Y}$ is then given by $\Delta\sqrt{(w - w_0)}/\Delta w$ times the mean value of the function $2\sqrt{(w - w_0)}.\ Y\,\partial\mu/\partial Y$ over the interval concerned. However, in non-equatorial regions the values of μ' for the extraordinary ray may be obtained directly in a form which may be integrated, by adding a small correcting term to the longitudinal approximation.

Writing $C = 1 - \cos\theta$ and retaining only powers of C up to the second, equation (10) gives for the extraordinary ray
$$\mu = t(1 + AC + ABC^2) \tag{11}$$

where
$$A = \frac{X}{2(1 - X)(1 - Y)}, \qquad B = \frac{Y(1 - 3t^2)}{4(1 - X)}$$

and
$$t = \sqrt{\left(1 - \frac{X}{1 - Y}\right)}$$

is the longitudinal form of μ. Applying (8) to this relation, and writing $Y/2(1 - Y) = K$, gives

$$\mu' t = 1 + \frac{KX}{1 - Y} + UC + VC^2 \tag{12}$$

where

$$U = A\left\{1 + K - (1 + 3K)t^2 - \frac{2t^2}{1 - X}\right\}$$

and

$$V = BU - At^2\left\{\frac{1 + X}{1 - X} B + 3(1 + K)A\right\}$$

The variation of U and V with t, for various values of Y, is shown in Fig. 3.

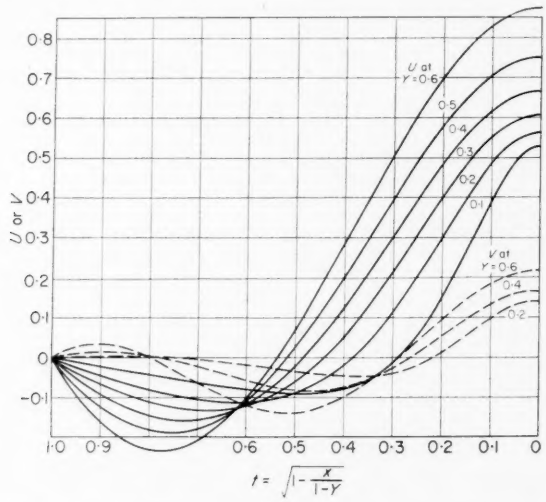


Fig. 3. Corrections to the longitudinal form of the group refractive index for the extraordinary ray.

$$\mu' \sqrt{\left(1 - \frac{X}{1 - Y}\right)} \fallingdotseq 1 + \frac{XY}{2(1 - Y)^2} + U(1 - \sin I) + V(1 - \sin I)^2$$

The correction $UC + VC^2$ is always largest at reflection ($t = 0$). The exact expression for μ' is then

$$\mu' t = \frac{2 - Y}{(1 - Y)\sqrt{(2 + 2\cos^2\theta)}} \tag{13}$$

The expansion of this expression in powers of C gives the correction $UC + VC^2$. For south-east England (dip = 67°), $C = 0{\cdot}08$ and the corrections UC and VC^2 amount to 4 and 0·08 per cent, respectively, at $t = 0$. The exact correction, from equation (13), is 4·076 per cent. A graph of $D = UC + VC^2$ is thus sufficiently accurate for determining μ' to about five significant figures.

At other locations, the correction $UC + VC^2$ yields values of $\mu' t$ accurate to within 0·2 per cent for I greater than 45°. Use of the single correction UC gives

VOL. 17 1959/6

values within 1 per cent for I greater than 30°, and within 6 per cent for all values of I.

The value of $\bar{\mu}'$ given by equation (12) for linear segments ($dh \propto dN$) is, writing $X/(1-Y) = X'$,

$$\bar{\mu}' = \frac{1}{\Delta X'} \int \mu' dX' = \frac{2}{\Delta X'} \int \mu' t \, dt = 2(1 + K + \bar{D}) \frac{\Delta t}{\Delta X'} - \frac{2}{3} K \frac{\Delta(t^3)}{\Delta X'}$$

where $\bar{D}$ is the mean value of D and may be estimated from Fig. 3.

Acknowledgements—This work was begun during the tenure of a post-graduate scholarship in science awarded by the University of New Zealand, and continued with the help of an overseas scholarship awarded by the Commissioners for the Exhibition of 1851. I would like to thank Dr. J. O. Thomas for his help with the preparation of this paper.

References

Budden K. G.	1954	*Conference on the Ionosphere* p. 332. Physical Society, London.
Jackson J. E.	1956	*J. Geophys. Res.* **61**, 107.
Kelso J. M.	1954	*J. Atmosph. Terr. Phys.* **5**, 11.
King G. A. M.	1954	*J. Atmosph. Terr. Phys.* **5**, 245.
Schmerling E. R.	1958	*J. Atmosph. Terr. Phys.* **12**, 8.
Thomas J. O.	1959	*Proc. Inst. Radio Engrs., N.Y.* **47**, 162.
Thomas J. O., Haselgrove J. and Robbins A. R.	1958	*J. Atmosph. Terr. Phys.* **12**, 46.
Titheridge J. E.	1959	*J. Atmosph. Terr. Phys.* This issue p. 110.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 110 to 125. Pergamon Press Ltd. Printed in Northern Ireland

The use of the extraordinary ray in the analysis of ionospheric records

J. E. TITHERIDGE
Cavendish Laboratory, Cambridge

(*Received* 29 *April* 1959)

Abstract—When the ordinary ray trace on an $h'(f)$ record is used to compute an electron density profile ($N(h)$ curve) assumptions have to be made about (a) the form which the $h'(f)$ curve would have taken at frequencies less than those actually employed, and (b) whether or not there is appreciable ionization in the "valley" between the E- and F-layers. In this paper it is shown how, by considering both the ordinary and extraordinary ray traces on the $h'(f)$ record, it is possible to avoid both these assumptions to a considerable extent, and to deduce something about the electron distribution in the lower ionosphere and in any valley.

The method is applied to some experimental records, and it is shown (a) that neglect of the low-lying ionization leads to an overestimate of the height of the night-time F-region ionization of about 20 km where the plasma frequency is 2 Mc/s, and up to 15 km near the peak of the layer, and (b) that the "valley" between the E- and F-layers is small and nearly "full".

1. INTRODUCTION

RECENTLY several methods have been used for deducing the distribution of electron density (N) as a function of height (h), as given by an $N(h)$ curve, from observations of $h'(f)$ curves which show the observed equivalent or virtual height (h') as a function of the radio-frequency (f). In the latest forms of these calculations it has been usual to allow for the effect of the earth's magnetic field, and to consider only the ordinary ray trace of the $h'(f)$ curve. Two assumptions have to be made in the analysis. (a) The form of the $h'(f)$ curve has to be assumed for frequencies below the lowest frequency of observation, often about 1 Mc/s. This requires, in effect, some assumption about the amount and distribution of the "underlying" or "low-level" ionization for which the plasma frequency is less than the lowest frequency of observation. (b) If the experimental $h'(f)$ curve shows sharp maxima at one or more values of f it could correspond either to a vertical tangent on the $N(h)$ curve, or to a re-entrant curve of the type shown in Fig. 1(a, b). When a curve is re-entrant we shall use the common nomenclature and speak of a "valley" between the two electron peaks. The size and shape of this valley must be assumed before the $N(h)$ curve can be calculated for the upper layer.

The simplest assumptions, and those usually made, are that the $h'(f)$ curve would have been horizontal at frequencies less than the lowest observed frequency, and that there are no "valleys". The inadequacy of these assumptions can lead to considerable errors in the calculated $N(h)$ curves.

It is the purpose of this paper to show that when the $h'(f)$ curves are available for both the ordinary and extraordinary rays the necessity for making these assumptions is to some extent removed. The analysis can also provide useful information about the amount of low-level ionization.

A simple, approximate relation is derived in Section 2. This gives the total amount of low-level ionization directly in terms of the difference between the

VOL.
17
1959/6

virtual heights of the ordinary and extraordinary rays at a single pair of frequencies. The coefficients in this relation are tabulated for different latitudes. The procedure is basically the same as that described by WATTS (1958) who used a low-frequency sounder to record the extraordinary ray *below* the gyrofrequency. Both methods require that the plasma frequency of the low-level ionization is always much less than the wavefrequency used, although this requirement is less rigid in the present work which makes allowance for the retardation of the ordinary ray. The results are similar to those obtained from the procedure described by JACKSON (1956)

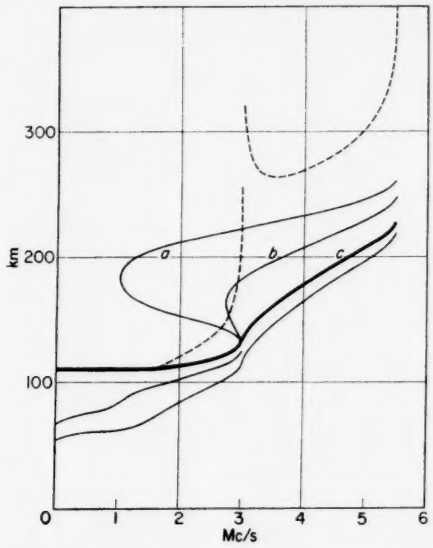


Fig. 1. The broken line shows a typical noon $h'(f)$ curve. The solid lines give some of the infinite number of possible computed height curves. The thick line is the result of the normal analysis.

in which an $N(h)$ curve is recalculated with different sized blocks of low-level ionization, until a curve is obtained which gives the correct extraordinary ray virtual heights.

For accurate calculation of the amount of low-level ionization, and of the errors in the normal $N(h)$ curves, it is necessary to consider not only the total amount of low level ionization, but also how it is distributed. In Section 3 it is shown that a consideration of this kind leads to the following procedure. An ordinary ray $h'(f)$ curve is analysed by any standard method. The resultant $N(h)$ curve is used to calculate the extraordinary ray virtual heights at a number of frequencies. The differences Δ between these heights and those experimentally observed are then plotted against a predetermined function of frequency. This plot defines a straight line whose intercept gives the total amount of ionization which must be added below the original $N(h)$ curve, and whose slope gives a measure of the distribution of this ionization. A simple function of these two parameters is then used to correct the $N(h)$ curve previously deduced from the ordinary $h'(f)$ trace alone.

The analysis is particularly useful for obtaining information about the amount of ionization below the night-time F-layer. The inherent accuracy of the method is about 1 per cent, so that the accuracy of the results depends almost entirely on the accuracy with which the virtual heights can be measured. The individual values of Δ are typically about 30 km, and an over-all accuracy of about 10 per cent is normally possible.

The same general procedure is used to determine the size of the valley between the E- and F-layers. Records taken at Slough near mid-day indicate a mean valley width of about 10 km. The heights obtained from the normal monotonic $N(h)$ analysis are then about 6 km and 4 km too low in the *F1*- and *F2*-regions.

When calculating the size of a valley it is important that the method used for obtaining the virtual height of the extraordinary ray from a calculated $N(h)$ curve should correspond exactly to that used in obtaining this curve. The reason for this is that differences in the frequencies at which the curves are read, and in the assumed shape of the $N(h)$ curve between the measured points, can cause errors of several kilometres in the calculated virtual heights at frequencies which just penetrate the layer. The method used in the present work is described in another paper (Titheridge, 1959). It permits the accurate calculation of a real or virtual height curve in about 10 min, using the same coefficients. This method also simplifies the calculations required in determining the amount of low-level ionization, since the calculation of an $N(h)$ curve and of the corresponding extraordinary ray virtual height can be carried out in one short operation.

The effect of electron collisions is neglected throughout. The errors due to this have been investigated and are quite negligible at the frequencies (greater than 1 Mc/s) used in the present work.

VOL.
17
1959/6

2. The Approximate Amount of Low-density Ionization below an Ionospheric Layer

The apparent or virtual height at which a radio wave is reflected in the ionosphere is given by

$$h' = \int_0^{h_0} \mu' \, dh \tag{1}$$

where h_0 is the actual height of reflection and μ' is the group refractive index, a function of the wavefrequency f and the electron density N.

An extraordinary wave of frequency f_x has the same actual height of reflection as the ordinary wave of frequency f_0 when

$$f_0^2 = f_x^2 - f_x f_H \tag{2}$$

The difference in the virtual heights of the two rays is then given by

$$h_x' - h_0' = \int_0^{h_0} (\mu_x' - \mu_0') \, dh \tag{3}$$

where μ_x' and μ_0' are the group refractive indices for the extraordinary and ordinary rays at the "corresponding" frequencies f_x and f_0.

For vertical propagation with a fixed magnetic dip angle I, μ_x' and μ_0' are functions of the two parameters $Y_0 = f_H/f_0$ and $X_0 = Ne^2/\pi m f_0^2 = f_N^2/f_0^2$ where

f_H is the gyrofrequency and f_N is the plasma frequency. The values of μ_x' and μ_0' increase almost linearly with X_0 when X_0 is small (Fig. 2), while for a fixed value of X_0 the difference $\mu_x' - \mu_0'$ is found to be roughly proportional to Y_0. Thus a constant K can be found such that

$$\mu_x' - \mu_0' = K \frac{N}{{f_0}^2} \cdot \frac{f_H}{f_0} \tag{4}$$

to within about 5 per cent for $f_N < \frac{1}{2} f_0$ and $f_0 > 1$ Mc/s. Inserting this relation in equation (3), we get

$$\int_0^{h_0} N \, dh = \frac{{f_0}^3}{K f_H} (h_x' - h_0') \tag{5}$$

This simple relation may be used to calculate the total amount of ionization below an ionospheric layer from the observed virtual heights of the ordinary and extraordinary rays at the two corresponding frequencies f_0 and f_x. The value of K was calculated for different values of the magnetic dip angle I, and is tabulated against sin I in Table 1. With these values equation (5) gives the total amount of low-level ionization with an accuracy of about 5 per cent.

Table 1. The coefficients K and K' in equations (5) and (6) (valid for $f_N < \frac{1}{2} f_0$ and $f_0 > 1$ Mc/s)

Sin I	0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
$K \cdot 10^5$	0	0·37	0·95	1·72	2·67	3·76	4·98	6·26	7·58	8·90	10·22
K'	—	—	0·26	0·44	0·68	1·00	1·39	1·84	2·34	2·88	3·46

Calculations of the ratio $(\mu_0' - 1)/(\mu_x' - \mu_0')$ show that it is independent of N and given by $f_0/K' f_H - 1/5$ where K' depends on I and is given in Table 1. Thus we get

$$\frac{h_0' - h_0}{h_x' - h_0'} = \frac{\int(\mu_0' - 1)\, dh}{\int(\mu_x' - \mu_0')\, dh} = \frac{f_0}{K' f_H} - \frac{1}{5}$$

and the actual height of reflection at the frequency f_0 is given by

$$h_0 = h_0' - \left(\frac{f_0}{K' f_H} - \frac{1}{5} \right) (h_x' - h_0') \tag{6}$$

The tabulated values of K' are accurate to within $\pm 0{\cdot}1$ for $f_N < \frac{1}{2} f_0$ and $f_H < f_0 < 4 f_H$.

Near the equator, $h_x' - h_0'$ is too small to measure accurately. At other latitudes equation (5) provides a direct estimate of the total amount of low-level ionization, while equation (6) gives the group-retardation caused by this ionization. Both results assume that the retardation is caused mainly by ionization with a plasma frequency of less than half the ordinary ray frequency f_0. However even when this condition is not fulfilled, equation (5) may still be applicable with reasonable accuracy. Thus if the electron density is supposed to increase linearly

with height from $f_N = \frac{1}{2}f_0$ to $f_N = f_0$, the value of $(h_x' - h_0')/\int N\,dh$ for this region (for conditions in south-east England) differs from the value given by equation (5) by only +29, −7 and −29 per cent at $f_0 = 1{\cdot}2$, 1·6 and 2 Mc/s, respectively. So for f_0 between about 1·3 and 1·8 Mc/s, equation (5) may be used to determine the total amount of ionization with $f_N < f_0$ to within about 20 per cent provided that the ionization gradient is fairly constant over the region in which $\frac{1}{2}f_0 < f_N < f_0$.

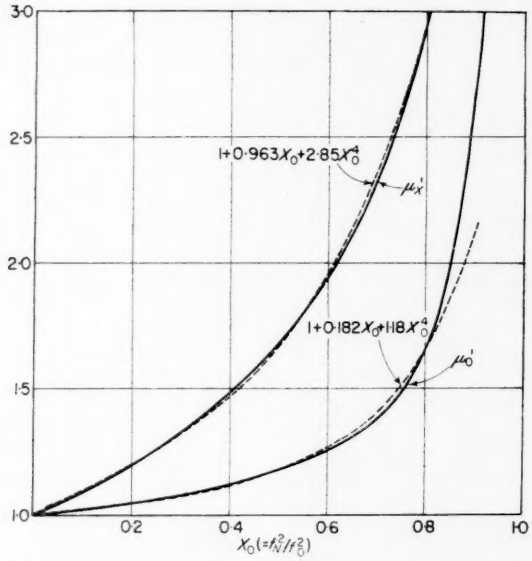


Fig. 2. The solid lines show the group refractive indices for the ordinary and extraordinary rays (in south-east England) at the two corresponding frequencies $f_0 = 2{\cdot}0$ Mc/s and $f_x = 2{\cdot}68$ Mc/s. The broken lines give the polynomial approximations used for $X_0 < 0{\cdot}8$.

3. Calculation of the Amount and Approximate Distribution of Low-level Ionization

3.1. *The relations used*

The previous section uses an approximation of the form $\mu' = 1 + k\,.\,X_0$ for the group refractive index μ'. This is accurate only for X_0 less than about 0·25, so that the final results are strictly valid only when the plasma frequency f_N of the low-level ionization is always less than half the ordinary wavefrequency. In general, f_N will vary continuously from 0 to $f_{\min}$, where $f_{\min}$ is the lowest frequency at which the ordinary ray $h'(f)$ curve is measured.

An approximation of the form

$$\mu' = 1 + K_1X_0 + K_2X_0^{\,r} \tag{7}$$

(where r is typically 3 or 4) is accurate to within a few per cent for values of X_0 up to about 0·8, i.e. for f_N up to $0{\cdot}9\,f_0$ (Fig. 2). Equation (7) may thus be used for the region in which $f_N < f_{\min} - 0{\cdot}2$ Mc/s, provided $f_{\min} < 2$ Mc/s. For the region in which f_N is between $f_{\min} - 0{\cdot}2$ and $f_{\min}$ the electron density will be assumed to increase linearly with height. Let the thickness of this second region be D km.

VOL. 17 1959/6

Then if $h_0'(f_0)$ and $h_x'(f_x)$ are the observed virtual heights of the ordinary and extraordinary rays, and $h_1(f_N)$ is the computed height curve obtained from h_0' on the assumption that there is no low-level ionization, it is shown in the Appendix that

$$\Delta = h_x'(f_x) - h_x''(f_x) = a\int N\,dh + b\int N^r\,dh + cD \tag{8}$$

where the integrals are taken over the region in which $f_N < f_{\min} - 0{\cdot}2$ Mc/s, and $h_x''(f_x)$ is the extraordinary ray virtual height curve calculated from $h_1(f_N)$; a, b and c are known functions of frequency, and can be calculated by the method described in the Appendix.

Table 2. The coefficients in equations (8) and (9) for conditions in south-east England ($I = 67°, f_H = 1{\cdot}18$ Mc/s)

f_0	1	1·2	1·4	1·6	1·8	2	2·2	2·4	2·6	2·8	3
$a\,.\,10^4$	—	0·967	0·595	0·398	0·283	0·209	0·161	0·128	0·104	0·086	0·072
$b\,.\,10^{12}$	—	1·428	0·635	0·356	0·232	0·170	0·135	0·112	0·095	0·082	0·069
c	—	1·03	1·40	1·28	1·17	1·05	0·93	0·81	0·70	0·61	0·58
$\alpha\,10^4$	0·157	0·157	0·149	0·141	0·134	0·125	0·118	0·110	0·104	0·099	0·093
$\beta\,10^{12}$	0·50	0·50	0·42	0·38	0·34	0·30	0·27	0·25	0·23	0·21	0·19
γ	5·26	4·26	3·58	3·08	2·68	2·36	2·10	1·90	1·74	1·60	1·48

Δ is seen to depend only on the amount and distribution of the low-level ionization, and from the values of Δ at three different frequencies the values of $\int N\,dh$, $\int N^r\,dh$ and D can be calculated. The errors in the computed height curve $h_1(f_N)$ also depend only on the low-level ionization (Appendix a) so that the corrected heights are given by

$$h(f_N) = h_1(f_N) - (\alpha\int N\,dh + \beta\int N^r\,dh + \gamma D) \tag{9}$$

The values of a, b, c, α, β and γ for conditions in south-east England (with $r = 3$ and $f_{\min} = 1{\cdot}2$ Mc/s) are given in Table 2. Their calculation is described in Appendix (b). Using these coefficients, the values of Δ given by equation (8) are accurate to about 1 per cent.

3.2. *The method of calculation*

Equation (8) (with $r = 3$) may be rearranged to give

$$\frac{\Delta - cD}{a} = \int N\,dh + \frac{b}{a}\int N^3\,dh \tag{10}$$

Using this relation the steps in the analysis of a virtual height record are as follows:

(i) The ordinary ray trace of an $h'(f)$ record is analysed by the standard method assuming that there is no ionization with f_N less than the lowest

recorded frequency $f_{\min}$. This gives the normal computed height curve $h_1(f_N)$ (the upper solid line in Fig. 5).

(ii) The extraordinary ray virtual heights corresponding to this curve are calculated, and subtracted from the observed heights to give the value of Δ at a number of frequencies (Fig. 5).

(iii) $(\Delta - cD)/a$ is plotted against b/a using (for south-east England) the coefficients a, b/a and c obtained from Table 2. The value of D (which is generally small) is adjusted until the plotted points lie near a straight line. The intercept of this line gives $\int N\,dh$, and the slope gives $\int N^3\,dh$.

(iv) The values of $\int N\,dh$, $\int N^3\,dh$ and D are inserted in equation (9) to give the final curve $h(f_N)$. This curve agrees exactly with the observed ordinary ray virtual heights, and normally agrees with the virtual heights of the extraordinary ray to well within the accuracy to which these can be measured. (Values of $\int N\,dh$, $\int N^3\,dh$ and D can of course be found so that the curve $h(f_N)$ agrees *exactly* with the observed extraordinary ray virtual heights at any three frequencies).

In the routine analysis of records of average accuracy, the value of $\int N^3\,dh$ cannot normally be determined with a worth-while accuracy. It is then preferable to assume some particular form for the distribution of the ionization with $f_N < f_{\min} - 0{\cdot}2$, so that only $\int N\,dh$ and D need be calculated. A linear increase of f_N with height from $f_N = 0$ to $f_N = f_{\min} - 0{\cdot}2$ provides a simple model, with a value of $\int N^3\,dh / \int N\,dh$ intermediate between the large values occurring near sunrise and sunset and the smaller values near midnight. For $f_{\min} = 1{\cdot}2$ Mc/s, this model gives $\int N^3\,dh = \frac{3}{7}\int N\,dh$ and equation (10) becomes

$$d\Delta = \int N\,dh + eD \tag{11}$$

where

$$\frac{1}{d} = a + \frac{3}{7}b \quad \text{and} \quad e = cd$$

Step (iii) above then consists simply of plotting $d\Delta$ against e. The intercept of the best straight line through the plotted points gives $\int N\,dh$, while the slope gives D. The corrected heights are obtained as before from equation (9), which now becomes

$$h = h_1 - \left(\alpha + \frac{3}{7}\beta\right)\int N\,dh - \gamma D \tag{12}$$

4. Applications

4.1. *Approximate calculations using the virtual heights measured at a single pair of frequencies*

Fig. 3 shows the results obtained from some night-time $h'(f)$ records made at Slough. The ordinary ray virtual height (h_0') recorded at $f_0 = 1{\cdot}4$ Mc/s was subtracted from the extraordinary ray height (h_x') at the corresponding frequency of 2·11 Mc/s. The difference of about 30 km could be obtained to within 3 to 5 km by using the second- or third-order reflections. Assuming the underlying ionization

VOL.
17
1959/6

to have a constant plasma frequency $f_N = 0{\cdot}6$ Mc/s ($N = 4500$ els/cm^3) and a thickness b km, b was found from equation (5). The circles in Fig. 3(a) show the average of the values of b obtained on seven of the international quiet days in December 1953. The amount of the underlying ionization is seen to be quite large, and almost constant throughout most of the night. The upper and lower crosses at each hour give the average of all values greater or less than the over-all mean, and show how the day-to-day variations are smoothed out as the night passes.

The distribution of ionization obtained at 0000 is shown in Fig. 3(b). It differs considerably from Fig. 5, which gives the result of a full analysis of the same records.

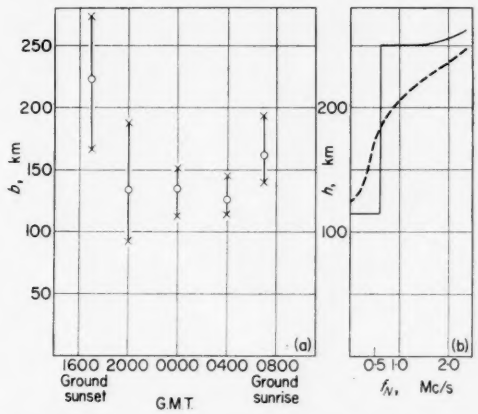


Fig. 3. The amount of low-level ionization ($f_N < 1{\cdot}4$ Mc/s) at Slough during December 1953. (a) The thickness of ionization, assuming f_N is constant and equal to 0·6 Mc/s, as obtained from $h_x' - h_0'$ at $f_0 = 1{\cdot}4$ Mc/s. The circles show the mean results from seven quiet days; the crosses are the average of all values greater or less than the over-all mean. (b) The solid line shows the assumed distribution of ionization at 0000. The broken line is the more accurate result of Fig. 5.

The value of $\int N\,dh$ obtained from Fig. 5 for $f_N < 1{\cdot}4$ Mc/s is, however, only 15 per cent less than that given in Fig. 3. A simple application of equation (5) thus gives, as expected, reasonably accurate estimates of the total amount of ionization with $f_N < f_0$. It does not give any idea of the distribution of this ionization and it does not give the amount of ionization below some known height. For when, as in the example considered, there is a large amount of ionization with f_N between $\frac{1}{2}f_0$ and f_0, equation (6) will seriously underestimate the required real height correction at $f_N = f_0$.

Equation (6) does however apply to reflections from the sharply bounded "sporadic-E" layers, which show little or no group retardation. Only records showing a dense sporadic-E layer with a definite penetration frequency can be used, since otherwise the ordinary and extraordinary ray traces cannot be separated. The value of $h_x' - h_0'$ then gives the total amount of ionization below the layer, the true height of which can be found from equation (6).

Fig. 4(b) shows the values of $\int N\,dh$ calculated from night-time $h'(f)$ records taken at Slough during June 1954. Each point is the mean of from 3 to 5 days, the standard deviations being shown by the vertical lines. The corresponding

heights of reflection are shown in Fig. 4(a). The records at 2000 and 0400 (40 min before ground sunset and after ground sunrise) could only be measured at frequencies above 1·4 Mc/s, and the small E-layer was not visible although its effect is unmistakable in the figure. The small pre-sunrise increase to 0300 occurred on all days and is probably significant.

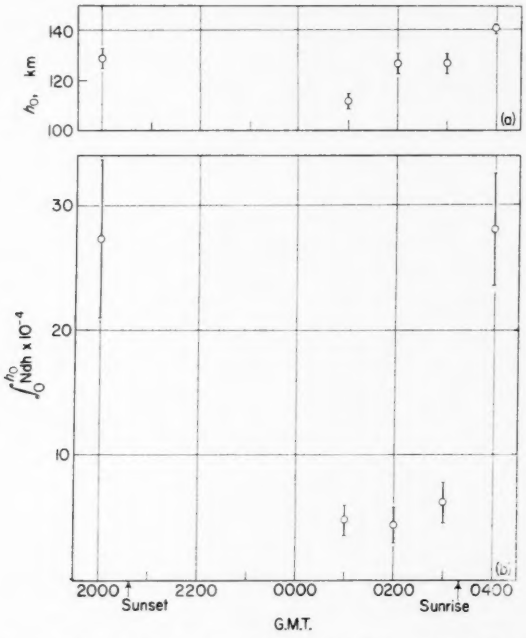


Fig. 4. The amount of ionization below a known height, as obtained from E_s-reflection at Slough during June 1954. (a) The real height of the E_s-layer. (b) The total amount of ionization below the layer. The vertical lines show the standard deviation in the day-to-day variations.

4.2. *The full analysis*

Fig. 5 shows the various steps in the calculation of a complete $h(f_N)$ curve. The broken lines give the average of the $h'(f)$ curves recorded at 0000 on seven of the international quiet days in December 1953. The upper solid line shows the curve $h_1(f_N)$ obtained by analysing the ordinary ray curve h_0', assuming this curve to be parallel to the frequency axis for $f \leqslant 1{\cdot}2$ Mc/s. The dotted line h_x'' gives the extraordinary ray virtual heights calculated from h_1. These heights are from 10 to 30 km below the observed heights h_x', giving from equation (11) the values $h(1{\cdot}2 \text{ Mc/s}) - h(1 \text{ Mc/s}) = D = (9 \pm 1)$ km, and $\int N\,dh = (26 \pm 5) \times 10^4$ km els/cm^3.

The lower solid line gives the final curve (h) obtained from equation (12). The extraordinary ray virtual heights calculated from this curve are shown by the circles in Fig. 5, and are within 1 km of the observed heights. The final curve h thus agrees exactly with the observed values of h_0' and differs from the observed values of h_x' by less than the experimental error. The shape of this curve at frequencies below 1 Mc/s is not significant, and any alternative shape (such as

that shown dashed) is equally valid provided that it has the same value of $\int N\,dh$. The total electron content of the final curve (i.e. the value of $\int N\,dh$ taken up to the peak of the layer) is 43 per cent greater than for the uncorrected curve h_1.

Some similar results from records taken at Slough during June 1957 are shown in Fig. 6. Each curve is the average over three of the international quiet days. The crosses show the lowest frequency (f') at which the virtual height of the

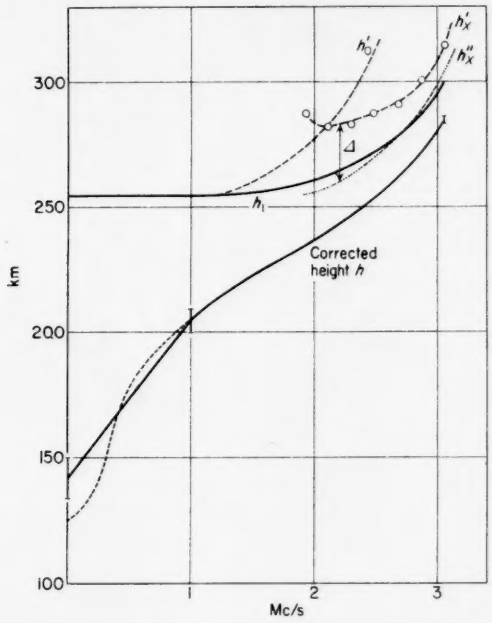


Fig. 5. The mean result from the analysis of records taken at Slough at 0000 GMT on seven of the international quiet days in December 1953. The broken lines show the observed $h'(f)$ curves. The upper solid line gives the heights computed from h_0', and the dotted line (h_x'') shows the corresponding extraordinary ray virtual heights. The lower solid line gives the corrected heights, the corresponding values of h_x' being shown by the circles. The vertical lines indicate the possible error in the calculated heights and in the amount of low-level ionization.

ordinary ray could be measured, and the height and gradient at this point are the lowest fixed features of the curves. Possible forms for the actual variation at $f_N < f'$ are shown dashed, maintaining the same value of $\int N\,dh$ as the solid curves.

The true shape of the lower ionosphere could be more definitely determined if accurate measurements were made on a larger number of records, when the value of $\int N^3\,dh$ could also be obtained. The curves shown do, however, give a good idea of the state of the lower ionosphere, and show very clearly the gradual decay of the E-region ionization after sunset and the large increase before sunrise, and the considerable errors in the normal analysis at these times. The curves of Fig. 6 will slightly underestimate the average amount of the low-level ionization, since the records with the clearest low-frequency traces were chosen and these tended to

be those in which $h_x' - h_0'$ was least. The day-to-day variations in the calculated values of $\int N\,dh$ are, however, generally less than ± 20 per cent.

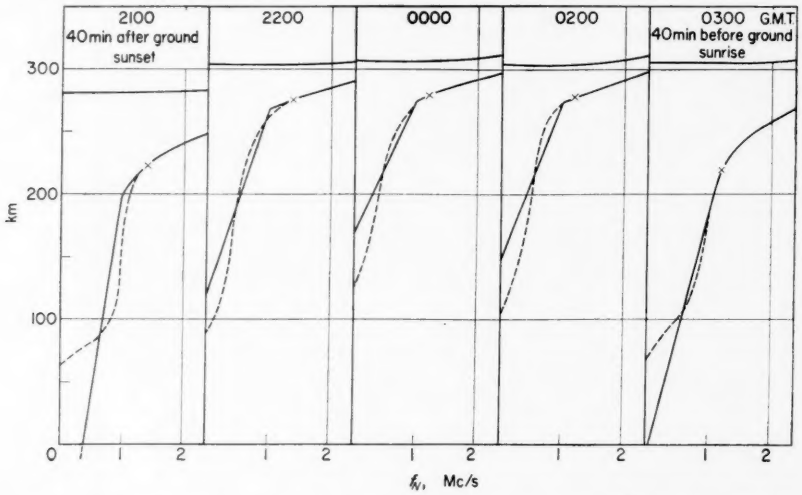


Fig. 6. The lower ionosphere at night. Slough, sunspot minimum (June 1957). The upper lines give the result of the normal analysis. The lower lines are the result of the present analysis. The broken lines give a more likely distribution of ionization, maintaining the same value of $\int N\,dh$.

VOL.
17
1959/6

5. VALLEY IONIZATION

5.1. *Calculation of the size and shape of the valley between two ionospheric layers*

The normal type of $N(h)$ analysis assumes that the electron density always increases with height. When there is a large cusp on the $h'(f)$ curve this gives a real height curve of the type shown in Fig. 1(c), with a near-vertical tangent. This is not, however, the only possible solution of the $h'(f)$ curve, since the $N(h)$ curve may not be monotonic but could have a "valley" or region in which the electron density is less than the value at the peak of the lower layer. Thus in Fig. 1 the curves (*a*), (*b*) and (*c*) all give the same ordinary ray $h'(f)$ curve, and by considering this curve alone it is not possible to determine the correct real height curve.

The extraordinary ray virtual heights corresponding to the curve (a) in Fig. 1 are, however, greater than the heights corresponding to the curve (c), due to the greater amount of low-density ionization. So the observed virtual heights of the extraordinary ray can be used to determine which of the possible solutions of the ordinary ray curve is the correct one.

The size and shape of any valley between two layers can be determined by the same procedure as was used earlier to determine the amount of low-level ionization. The ordinary ray $h'(f)$ curve is first used to calculate the normal, monotonic $N(h)$ curve. The extraordinary ray virtual heights corresponding to this curve are then calculated and subtracted from the observed heights to give Δ. Then just as the values of Δ calculated at frequencies reflected from the E-layer tell us how much ionization must be added below this layer, so the values of Δ at frequencies greater

than the critical frequency f_cE of the E-layer tell us how much ionization must be added below the F-layer. If the shape of the $N(h)$ curve has previously been accurately determined up to the peak of the E-layer, this additional ionization can only be put in a valley.

Equation (8) is used as before to determine the values of $\int N\,dh$ and $\int N^3\,dh$, which give the amount of ionization to go in the valley. D gives the amount by which the increase in height from $f_N = f_cE$ to $f_N = f_cE + 0{\cdot}2$ Mc/s must be changed. It is large and negative when a valley is present, since the ionization gradient at the base of the F-layer is considerably increased (Fig. 1). The corrected F-layer heights h given by equation (9) are thus greater than the heights h_1 calculated on the assumption of a monotonic $N(h)$ curve, and the amount by which the base of the F-layer must be raised gives the width of the valley.

For typical day-time records taken at Slough during sunspot minimum years, the values of Δ calculated for the F-layer are less than 40 km. The valleys required are then only from 0 to 15 km wide and have a mean electron density of at least 80 per cent of that at the F-layer maximum. These valleys are too small to determine accurately from normal $h'(f)$ records, for which the following approach is adequate.

Table 3. The effect of a 1 km wide valley in which the electron density increases linearly with height from $f_N = f_cE - 0{\cdot}2$ to $f_N = f_cE$ where f_cE is the critical frequency of the E-layer
(Angle of dip = 67°, f_H = 1·18 Mc/s)

(a) The increase in the virtual height of the extraordinary ray

$f_0 - f_cE$ =	0·2	0·4	0·6	0·8	1·0	1·2	1·4
f_cE = 1·6	1·10	0·47	0·27	0·12	0·04	−0·02	−0·03
2·0	1·22	0·58	0·33	0·17	0·09	0·04	0·01
2·4	1·33	0·67	0·38	0·21	0·12	0·07	0·04
2·8	1·44	0·74	0·43	0·25	0·16	0·11	0·07
3·2	1·53	0·79	0·46	0·27	0·19	0·13	0·08
3·6	1·60	0·83	0·48	0·28	0·20	0·14	0·08
4·0	1·65	0·86	0·50	0·30	0·21	0·14	0·08

(b) The increase in the true height of the F-region

$f_0 - f_cE$ =	0	0·4	0·8	1·2	2	3	4
f_cE = 1·6	1·0	0·66	0·50	0·39	0·28	0·20	0·15
2·4	1·0	0·68	0·52	0·41	0·30	0·22	0·17
3·2	1·0	0·69	0·53	0·43	0·32	0·24	0·19
4·0	1·0	0·69	0·53	0·44	0·32	0·24	0·19

5.2. *The approximate determination of the size of the valley between the E- and F-layers*

For a practical valley in which the electron density decreases only slightly, the values of Δ introduced when the valley is neglected are proportional to the

valley width and almost independent of the shape. For some assumed distribution of ionization in the valley, the increase in the true height of the F-layer and in the virtual height of the extraordinary ray can be calculated for a valley width of 1 km. The increases for a valley in which the electron density varies linearly with height from $f_N = f_c E - 0{\cdot}2$ to $f_N = f_c E$ are shown in Table 3, calculated for south-east

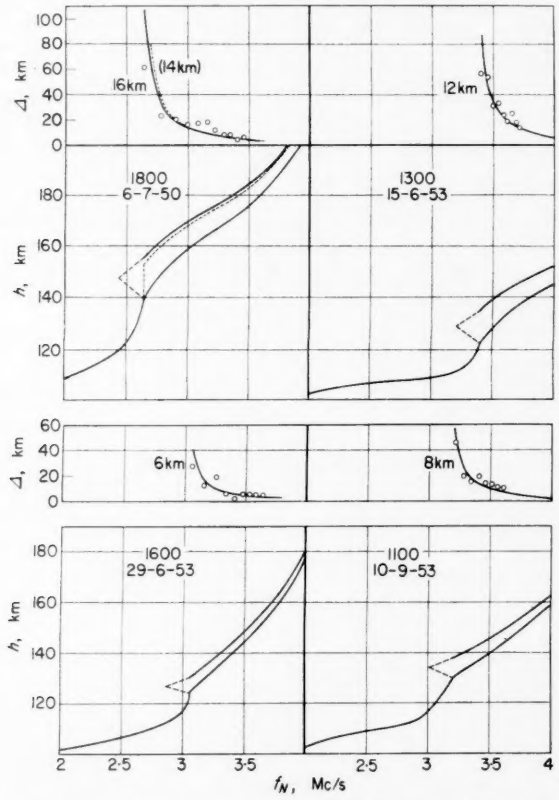


Fig. 7. The size of the valley between the E- and F-layers. The lower of the two $h(f_N)$ curves at each hour shows the result of the direct analysis of the ordinary ray trace of an $h'(f)$ record taken at Slough. The difference Δ between the observed virtual heights of the extraordinary ray and those calculated from this $h(f_N)$ curve are shown by the circles at the top of each figure. The curves through these points correspond to triangular valleys of the sizes indicated. The broken lines give the shape of the valley, and the solid lines give the corrected height of the F-layer. The dotted lines at 1800 give the result for a "full" valley.

England. From these figures the width of the valley, and thus the true height corrections to the F-layer, can be obtained from the amount by which the calculated virtual heights of the extraordinary ray must be increased to agree with the observed heights.

Some standard $h'(f)$ records, taken at Slough, were analysed in this manner. Only four of the available records showed clear extraordinary ray traces for both the E- and F-layers. The ordinary ray traces were analysed by the normal method,

and corrected for the low-level ionization. The resulting computed height curves are shown in Fig. 7 (the lower curve at each hour). The virtual heights of the extraordinary ray were calculated from these curves and subtracted from the observed heights to give the values of Δ plotted in Fig. 7 above the corresponding computed height curves; there is in all cases a sudden increase of about 40 km in Δ when $f > f_cE$. The curves drawn through the plotted points were calculated from Table 3(a) using the valley width which gave the best over-all fit. The widths required vary from 6 to 16 km. The agreement with the plotted values of Δ is probably as good as can be expected, since the virtual heights were measured from tracings with an accuracy of about 5 km, giving a possible error of 10 km in Δ. The final true height curves (obtained from Table 3b) are also shown in Fig. 7, with the assumed valley shape indicated by the dashed lines.

The calculations for the first record, which gave a valley width of 16 km, were repeated on the assumption of a perfectly full valley. The results are shown dotted; the width of the required valley is reduced to about 14 km. The close agreement between the two curves shows the comparatively small effect of the shape of the valley on the calculated heights of the F-region. A close approximation to the "full valley" distribution would of course be obtained from the direct ordinary ray analysis if a sufficiently small frequency interval was used; in the present work the interval was about 0·1 Mc/s.

OL.
17
9/60

Acknowledgements—This work was carried out during the tenure of an Overseas Scholarship awarded by the Commissioners for the Exhibition of 1851. I would like to thank Mr. J. A. Ratcliffe for his constant encouragement, and Dr. R. L. Smith-Rose (Director of Radio Research of the DSIR) for the loan of records from Slough. Thanks are also due to the Director of the Mathematical Laboratory in the University of Cambridge for permission to use the laboratories digital computor (EDSAC II) for calculating the values of μ' from which Tables 1, 2 and 3 were derived.

References

Budden K. G.	1954	*Conference on the Ionosphere*. p. 332. Physical Society, London.
Jackson J. E.	1956	*J. Geophys. Res.* **61**, 107.
Titheridge J. E.	1959	*J. Atmosph. Terr. Phys.* This issue p. 96
Watts J. M.	1958	*J. Geophys. Res.* **63**, 717.

Appendix

(a) *The derivation of equations* (8) *and* (9) *of Section* 3.1

Consider an $N(h)$ curve defined by the true heights of reflection h_T at a number of frequencies f_0. The height is supposed to be constant for values of the plasma-frequency f_N less than the lowest of these frequencies ($f_{\min}$). Then the corresponding series of virtual heights for the ordinary ray (h_0') is given by

$$h_T = M_0 h_0' \qquad \text{or} \qquad h_0' = M_0^{-1} h_T$$

where M_0 is the Budden matrix for the ordinary ray (denoted by L in the paper by Budden, 1954).

If there is some low-level ionization present (i.e. ionization with $f_N < f_{\min}$), the virtual heights become

$$h_0'(f_0) = M_0^{-1} h_T + R_0(f_0) \tag{1}$$

where R_0 is the group retardation caused by the low-level ionization. Similarly the virtual heights for the extraordinary ray (at the corresponding frequencies f_x) may be written

$$h_x'(f_x) = M_x^{-1} h_T + R_x(f_x) \tag{2}$$

If the ordinary ray virtual heights are analysed by the normal methods neglecting the low-level ionization, the computed heights h_1 are given by

$$\begin{aligned} h_1 &= M_0 h_0' \\ &= M_0 M_0^{-1} h_T + M_0 R_0 \\ &= h_T + M_0 R_0 \end{aligned} \tag{3}$$

The extraordinary ray virtual heights corresponding to these computed heights are given by

$$h_x''(f_x) = M_x^{-1} h_1 = M_x^{-1} h_T + M_x^{-1} M_0 R_0 \tag{4}$$

Subtracting (4) from (2), the difference between the actual values of the extraordinary ray virtual heights, and those obtained from the computed heights h_1, is given by

$$\Delta = h_x' - h_x'' = R_x - M_x^{-1} M_0 R_0 \tag{5}$$

The value of Δ thus depends only on the low-level ionization, and not at all on the heights of the reflecting layer.

When the plasma frequency f_N of the low-level ionization is less than $f_{\min} - 0{\cdot}2$, we may use the approximation

$$\mu' = 1 + a'N + b'N^r \tag{6}$$

where a' and b' are functions of frequency. The retardation due to this ionization is then $R = \int(\mu' - 1)\, dh = a'\int N\, dh + b'\int N^r\, dh$. If the electron density increases linearly with height in the region in which f_N is between $f_{\min} - 0{\cdot}2$ and $f_{\min}$, the retardation due to this region is simply proportional to its thickness D, and the full expression for R is

$$R = a'\int N\, dh + b'\int N^r\, dh + c'D \tag{7}$$

Using the subscripts 0 and x to denote the values of a', b' and c' for the ordinary and extraordinary rays, respectively, equation (5) becomes

$$\Delta = a\int N\, dh + b\int N^r\, dh + cD \tag{8}$$

where

$$a = a_x' - M_x^{-1} M_0 a_0'$$

$$b = b_x' - M_x^{-1} M_0 b_0'$$

and

$$c = c_x' - M_x^{-1} M_0 c_0'$$

VOL.
17
1959/6

The true heights h_T of the reflecting layer are obtained from equation (3) which becomes

$$h_T = h_1 - (\alpha \int N\, dh + \beta \int N^r\, dh + \gamma D) \tag{9}$$

where

$$\alpha = M_0 a_0', \qquad \beta = M_0 b_0' \qquad \text{and} \qquad \gamma = M_0 c_0'$$

(b) *Calculation of the coefficients in equations* (8) *and* (9)

The coefficients in equations (8) and (9) are given above in terms of the coefficients a' and b' of equation (6), which are determined empirically from calculated values of μ'. It is, however, simpler and more accurate to obtain the coefficients a, b and c directly from equation (8) using the values of Δ calculated from models with known values of $\int N\, dh$, $\int N^r\, dh$ and D.

Coefficient c is equal to the value of Δ caused by a step D of 1 km. Δ is calculated from equation (5), by the following steps:

(i) The ordinary and extraordinary ray retardations (R_0 and R_x), due to a step $D = 1$ km, are calculated at a number of frequencies.

(ii) $R_0(f_0)$ is treated as an observed $h'(f)$ curve and analysed to give a computed height curve.

(iii) This curve is used to calculate an extraordinary ray virtual height curve.

(iv) Subtracting this last curve from R_x gives the value of Δ as a function of frequency.

The result of step (ii) is the error ($M_0 R_0$ in equation 3) in the computed heights caused by the step D of 1km. It thus gives the value of γ to be used in equation (9).

The values of a, b and r were calculated from the values of Δ caused by blocks of ionization of unit thickness and with plasma frequencies of 0·4, 0·6 and 0·8 Mc/s. If, at some frequency, these values of Δ (calculated in the manner described above) are Δ_1, Δ_2 and Δ_3, respectively, then the value of r at that frequency is obtained from

$$0{\cdot}4^r(0{\cdot}6\Delta_3 - 0{\cdot}8\Delta_2) + 0{\cdot}6^r(0{\cdot}8\Delta, -0{\cdot}4\Delta_3) + 0{\cdot}8^r(0{\cdot}4\Delta_2 - 0{\cdot}6\Delta_1) = 0$$

For conditions in south-east England the value of r varies from 2·8 at $f_0 =$ from 1 to 1·4 Mc/s, to 3·3 at from 2 to 4 Mc/s. A mean value of $r = 3{\cdot}0$ was taken. For each pair of blocks the values of a and b are then calculated from equation (8) using the known values of $\int N\, dh$ and $\int N^3\, dh$. Averaging the three values of a and b obtained at each frequency gives the figures shown in Table 2. The values of α and β are calculated to fit the relation

$$M_0 R_0 = \alpha \int N\, dh + \beta \int N^3\, dh$$

Use of the tabulated coefficients a and b in equation (8) gives errors of less than 2 per cent for any value of f_N between 0·3 and 0·9 Mc/s. The relation is exact for f_N near 0·5 and 0·7 Mc/s. Equation (8) will thus be accurate to about 1 per cent for any normal distribution of ionization with f_N varying from 0 to 1 Mc/s.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 126 to 133. Pergamon Press Ltd. Printed in Northern Ireland

Ionization below the night-time *F*-layer

J. E. Titheridge
Cavendish Laboratory, Cambridge

(*Received* 29 *April* 1959)

Abstract—By making use of both the ordinary and extraordinary ray traces on $h'(f)$ records it is possible to estimate the amount and distribution of low-lying ionization, having plasma frequencies less than the lowest frequency recorded. By applying this method to $h'(f)$ curves obtained at night, it is possible to estimate the electron content of the *E*-region, even when its critical frequency is less than the lowest recorded frequency. Results are given for Slough and Watheroo for both summer and winter conditions, and for maximum and minimum sunspot numbers. Near midnight the amount of ionization below the *F*-region is equivalent to a constant density of about 4000 electrons/cm^3 extending down to a height of 130 km. The variation in the amount of this ionization near sunset gives a constant effective recombination coefficient of 2×10^{-8} cm^3/sec. *F*-layer heights calculated from the ordinary ray trace only are found to be too great by about 30 km at $f_N = 1$ Mc/s and 5 km at $f_N = 6$ Mc/s.

1. Introduction

Records showing the apparent height h' at which a radio wave is reflected in the ionosphere, as a function of the wavefrequency f, are taken hourly at many observatories. These records can be used to calculate the variation of electron density N with height h, and they provide the main source of information about the ionosphere. The direct analysis of an $h'(f)$ curve will, however, give the correct result only if there is no ionization with a plasma frequency $f_N = Ne^2/\pi m$ less than the lowest frequency ($f_{\min}$) at which the virtual heights were recorded.

It is shown in another paper (Titheridge, 1959) that this restriction can be largely overcome by using both the ordinary and the extraordinary ray traces of an $h'(f)$ record. In particular a method is given for determining both the total amount of ionization with $f_N < f_{\min}$, and the true heights of the reflecting layer. The present paper gives the results obtained from some night-time records taken at Slough (lat. 52°N, dip angle $I = 67°$) and at Watheroo (30°S, $I = 64°$) in both summer and winter and for maximum and minimum sunspot activity. The analysis gives the total amount of ionization present below the *F*-layer, so that the gradual decay of the *E*-region ionization after sunset, and the increase before sunrise, can be followed. The difference between the actual heights of the *F*-layer and those calculated by the normal ordinary ray analysis is also obtained.

2. The Method of Analysis

When an $h'(f)$ curve is available only for frequencies greater than $f_{\min}$ it is first supposed that the virtual height at all lower frequencies has the value appropriate to $f_{\min}$ and an $N(h)$ curve is calculated from the ordinary ray trace. Curves of this kind are shown by the broken lines in Figs. 1 and 2. This curve is then modified, in the manner described in Section 4.2 of Titheridge (1959), so as to provide the best over-all fit with the observed extraordinary ray virtual heights, while still retaining the correct virtual heights for the ordinary ray.

The records used in the present work could generally be used down to a frequency $f_{\min}$ of 1·2 Mc/s. The distribution of ionization with $f_N < 1{\cdot}2$ Mc/s is

then approximated to by two sections in each of which f_N increases linearly with height, from 0 to 1 Mc/s and from 1 to 1·2 Mc/s. The section from 1 to 1·2 Mc/s serves to give the correct gradient of ionization at the frequency $f_{\min}$. The lower section (from $f_N = 0$ to $f_N = 1$ Mc/s) gives the total amount of ionization with $f_N < 1$ Mc/s. The actual distribution of ionization in this region cannot normally be found since it does not have much effect on the observed $h'(f)$ curves. The height and gradient at the lowest observed frequency are thus the lowest fixed features of the final curves, and for $f_N < f_{\min}$ any alternative curves with the same total electron content will fit the observed $h'(f)$ curves almost equally well.

3. General Results

Fig. 1 shows the results obtained from night-time $h'(f)$ records made at Slough. The broken lines give the result of computations made from the extrapolated trace of the ordinary ray, and the full lines show the corrected curves. Each curve is the mean result over about 5 days. The records used were sufficiently accurate for the mean amount of low-density ionization to be determined to within about 10 per cent. The records showing the clearest low-frequency traces were used, and these tended to be the ones indicating the smallest amount of low-level ionization. The curves shown will thus slightly underestimate the average amount of this ionization. However, the day-to-day variations in this amount, for the records analysed, were generally less than ± 20 per cent. The day-to-day variations do not show any significant correlation with magnetic activity.

Except for records taken within $1\frac{1}{2}$ hr of sunrise or sunset, the size of the error made by using the extrapolated trace of the ordinary ray is seen to be fairly constant. Thus the height at $f_N = 2$ Mc/s is too great by 17 ± 4 km. The corresponding error at Watheroo (Fig. 2) is 16 ± 5 km. So the $N(h)$ curves as normally calculated must be lowered by about 30, 17, 11, 8 and 5 km at $f_N = 1$, 2, 3, 4 and 6 Mc/s, respectively.

The lower, linear, portions of the curves in Fig. 1 give the total amount of ionization with $f_N < 1$ Mc/s. The actual distribution of this "low-level" ionization is not known. The dotted lines in Fig. 1 indicate a possible distribution, maintaining the same total amount of ionization with f_N less than the lowest recorded frequency (marked by a cross) as the continuous line curves. The main features of the curves of Fig. 1 are summarized below.

(a) There are no significant seasonal changes. Except near 2200, the curves for summer and winter, at a given sunspot number, and at corresponding intervals throughout the night, are almost identical.

(b) The amount of low-level ionization at midnight, and in the early morning, is 70 per cent greater in sunspot maximum than in sunspot minimum. This increase can be explained entirely by the greater height of the F-layer if the low-level ionization is supposed to have a fixed mean electron density and a fixed lower limit of about 130 km. The ionization is equivalent in amount to a constant electron density $N = 3400$ els/cm^3 ($f_N = 0{\cdot}52$ Mc/s) extending from 130 km to the bottom of the F-layer (i.e. the height at $f_N = 1$ Mc/s, about 30 km below the minimum virtual height of the F-layer).

(c) The rate of decay of ionization after sunset appears to be about twice as

great at sunspot minimum as at sunspot maximum. Thus in sunspot minimum years the amount of low-level ionization has decreased to its approximately constant night-time value within 2 hr of ground sunset; at sunspot maximum the time taken is over 2 hr in summer and 4 hr in winter. The total electron

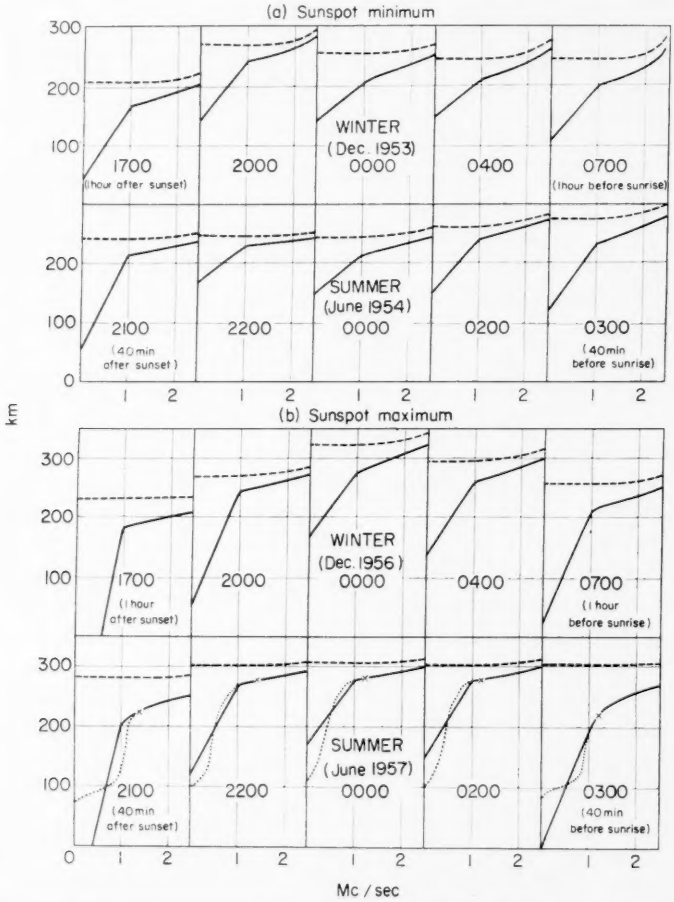


Fig. 1. The night-time ionosphere over Slough (lat. 52°N, dip angle 67°). Each curve is the average over about 5 days. The broken lines give the result of the direct analysis of the ordinary ray trace of an $h'(f)$ record. The solid lines show the result of correcting these curves so that they also agree with the extraordinary ray trace. The lower, linear portion of the solid curves serves only to give the total amount of low-density ionization. The dotted lines show a more likely form for the actual distribution of the ionization with f_N less than the lowest recorded frequency (marked by a cross).

density (below about 200 km) at 1 hr after ground sunset is 2·5 times greater in sunspot maximum than in sunspot minimum, although at sunset the difference is only about 50 per cent.

(d) At sunspot maximum, the total amount of ionization (below 220 km) at 1 hr before ground sunrise is eight times the corresponding amount near midnight. At sunspot minimum the increase is only about 40 per cent.

VOL. 17 1959/6

(e) The increase in ionization at 1 hr before ground sunrise is just as large in winter as in summer. Yet the winter records were taken 15 hr after ground sunset (and only 25 min after the sun had risen at a height of 100 km) while the summer records were only 7 hr after ground sunset (and 90 min after sunrise at 100 km).

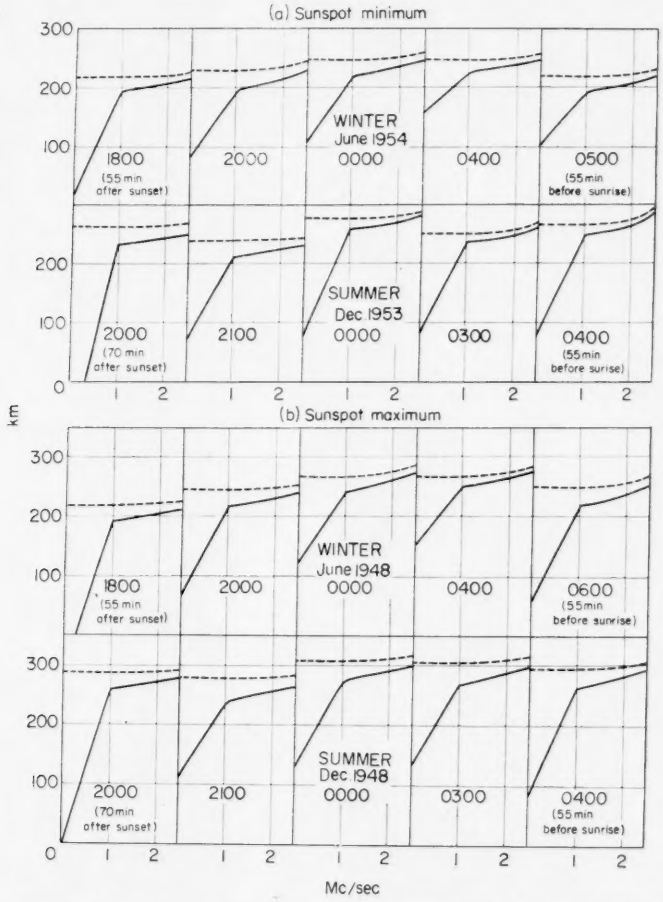


Fig. 2. The night-time ionosphere over Watheroo (lat. 32°S, dip angle 64°). The broken lines give the result of the normal ordinary ray analysis. The solid lines give the result of the present analysis. Each curve is the average over about 6 days.

The corresponding results for Watheroo are shown in Fig. 2. The day-to-day variations in the amount of low-level ionization were up to 2 : 1, so that the mean curves shown (from about 6 days) are slightly less accurate than the curves of Fig. 1. Near midnight the amount of low-level ionization is rather greater than at Slough, particularly at sunspot minimum. In December 1953 this ionization is equivalent to a constant electron density $N = 6000$ els/cm^3 ($f_N = 0{\cdot}7$ Mc/s) extending from 130 km to the bottom of the F-layer. All the other months are equivalent to $N = 4000$ els/cm^3 ($f_N = 0{\cdot}57$ Mc/s) from 130 km.

The pre-sunrise increase of ionization is again much less marked at sunspot minimum, and is at least as large in winter as in summer. The increase at 1 hr before ground sunrise is only one-quarter as large as at Slough, owing to the greater speed with which the sun rises nearer the equator. In contrast to the results at Slough, the rate of decay of the E-region ionization after sunset does not vary significantly with the sunspot number; the value of the apparent recombination coefficient is considered in the next section.

4. The Decay of Ionization after Sunset

Fig. 3 shows the variation in the E-region ionization throughout the winter night at Watheroo. The solid lines show the calculated curves, each curve being

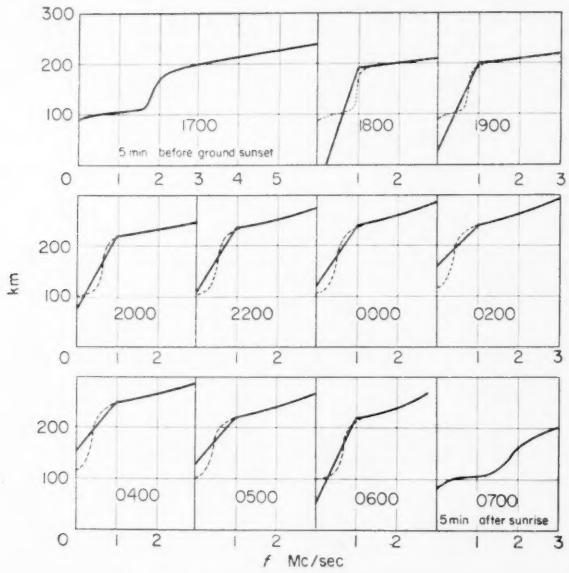


Fig. 3. The variation of the E-region ionization from sunset to sunrise at Watheroo during June 1948. Each curve is the average over 8 days. The solid lines are the calculated curves; the broken lines show a possible form for the actual variation of f_N with height.

the mean result from 8 days. The broken lines give a possible form for the actual variation of f_N with height, and show a smooth decay of ionization from sunset to 3 hr before sunrise. The amount of low-level ionization then begins to increase although the F-region ionization continues to decrease until about 40 min before ground sunrise.

In the first few hours after sunset, the height h_m of the bottom of the F-layer is fairly constant at about 200 km. It is then possible to determine quite accurately the total amount of ionization below this fixed height. The ratios of the total number of electrons $n(t)$ in unit column below 200 km at different times of night, to the corresponding number $n(0)$ at sunset, are shown by the dots in Fig. 4. The results for the three other months analysed at Watheroo (for the hours up to

VOL. 17 1959/6

midnight) are also shown; the more irregular variation in these months is largely due to the smaller number of days (about 5) which could be analysed to obtain the mean curves at each hour.

If the ionization is supposed to disappear after sunset by a recombination process, with a coefficient α, then $n(0)/n(t) = 1 + \alpha N_0 t$, where N_0 is the mean

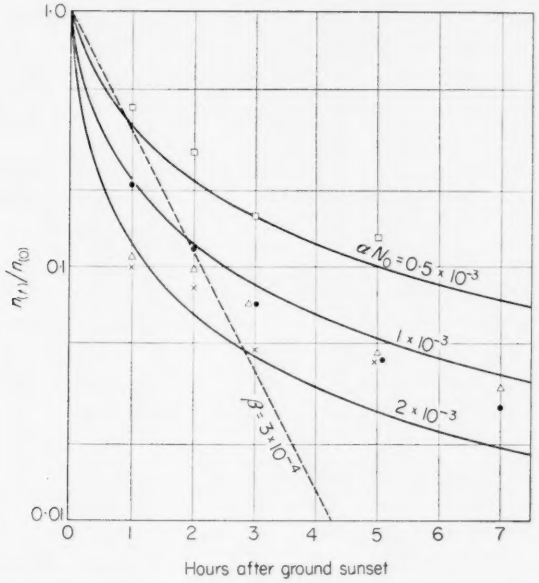


Fig. 4. The relative amount of ionization below the height h_m from ground sunset to midnight. $n(t)$ is the value of $\int_0^{h_m} N dh$ at t sec after ground sunset, and $n(0)$ is the value at sunset. The solid lines show the expected variation of $n(t)/n(0)$ for three different values of αN_0, where α is the effective recombination coefficient and N_0 is the mean electron density at sunset. The experimental points give the mean values of $n(t)n(0)$ over about 6 days. The values of αN_0 giving the best over-all fit for each month lead to the following results:

	h_m (km)	$\alpha N_0 \times 10^3$	$N_0 \times 10^4$	$\alpha \times 10^8$
● June 1948	200	1·17 ± 0·1	6·0	1·95 ± 0·2
× Dec 1948	250	1·65 ± 0·35	7·0	2·35 ± 0·5
△ June 1954	200	1·45 ± 0·3	8·5	1·7 ± 0·35
□ Dec 1953	220	0·42 ± 0·06	2·5	1·7 ± 0·2

electron density at sunset. The solid lines in Fig. 4 show the expected variation for three different values of αN_0. The broken line shows the variation which would be expected if the ionization disappeared by attachment with a coefficient $\beta = 3 \times 10^{-4}\ \text{sec}^{-1}$.

The experimental points are seen to follow curves of the recombination type. The values of αN_0 giving the best over-all fit are tabulated beneath the figure. The mean electron density N_0 is obtained by dividing the total amount of

ionization below the height h_m (at sunset) by $h_m - 100$. The resultant values of α for the four months do not differ significantly, and give a mean value

$$\bar{\alpha} = (1{\cdot}9 \pm 0{\cdot}2) \times 10^{-8} \text{ cm}^3/\text{sec}$$

For a region in which both α and N vary with height, calculations show that the total amount of ionization decays very nearly as if it had a constant recombination coefficient equal to the mean value of α, and a constant electron density equal to the mean value of N. Fig. 4 will thus give fairly accurately the mean value of α between about 100 and 200 km. The curves of Fig. 3 indicate that N is reasonably constant over this height range for the first 4 hr after sunset, so that α does not vary rapidly with height. The experimental points in Fig. 4 show the expected scatter of up to ± 20 per cent, but there is no systematic departure from the shape of the calculated curves. So the effective recombination coefficient does not change significantly (i.e. by more than about 30 per cent) between sunset and midnight.

This is in contrast to the results obtained by Mitra (1954) from low-frequency measurements, which required a value of α (at 100 km) decreasing from about 3×10^{-8} near sunset to 3×10^{-9} at midnight. The decrease was attributed to the presence of positive metallic ions, and these may only be important near 100 km. The present results agree with the day-time values of about $1{\cdot}5 \times 10^{-8}$ to 4×10^{-8} at the peak of the E-layer, obtained from eclipse measurements.

The above analysis cannot be applied at Slough, since the height of the F-layer is not sufficiently constant. A rough estimate of the mean effective recombination coefficient can however be made from the conclusion of Section 3(b) that the night-time ionization has a mean electron density of 3400 els/cm^3. This value is reached in about 4 hr at sunspot maximum and 2 hr at sunspot minimum, requiring a mean effective recombination coefficient of 2×10^{-8} and 4×10^{-8} cm^3/sec, respectively.

5. Conclusions

The general results obtained above lead to the following tentative picture of the night-time ionization at medium latitudes.

From the time of ground sunset, the ionization in the E- and F_1-regions (below about 200 km) decays with a constant effective recombination coefficient α of 2×10^{-8} cm^3/sec. The amount of ionization below the F-layer at midnight is equivalent to a constant density of about 4000 els/cm^3 extending down to 130 km. This ionization persists unchanged until a few hours before sunrise. There are no large seasonal or sunspot number variations in the amount of this ionization or in the value of α (except at Slough at sunspot minimum when the decay of ionization after sunset seemed abnormally rapid). The amount of ionization below 200 km starts to increase early and has increased significantly by 1 hr before ground sunrise. The amount of this increase does not vary with season, but changes from about 30 per cent in sunspot minimum to several hundred per cent at sunspot maximum.

Acknowledgements—This work was carried out during the tenure of an Overseas Scholarship awarded by the Commissioners for the Exhibition of 1851. Thanks

VOL.
17
1959/6

are due to Dr. R. L. SMITH-ROSE (Director of Radio Research of the D.S.I.R.) for the loan of records from Slough, to A. F. TILLOT (Observer in Charge, Watheroo Magnetic Observatory) for the loan of records from Watheroo, and to the Director of the Mathematical Laboratory in the University of Cambridge for permission to use the laboratory's digital computer (EDSAC II) to calculate the coefficients used in the analysis.

REFERENCES

MITRA A, P. 1954 Scientific Report No. 68. Pennsylvania State University.

TITHERIDGE J. E. 1959 *J. Atmosph. Terr. Phys.* **17**, 110.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 134 to 144. Pergamon Press Ltd. Printed in Northern Ireland

Sealed-room experiments on the equilibrium of ionization in air

O. C. Jones,* R. S. Maddever,† J. H. Sanders,
The Clarendon Laboratory, Oxford

(*Received* 4 *June* 1959)

Abstract—In an attempt to establish the correct form of the equation describing the equilibrium of ionization in air periodic measurements were made of Z and n, the concentrations of condensation nuclei and of small ions, inside a sealed room. It was found that n increased linearly with time in the initial stages, whilst Z decreased in agreement with an equation proposed by Nolan. These results did not fit the type of equilibrium equation usually assumed. In further experiments direct measurement showed that q, the rate of production of ion pairs, varied considerably with time, controlling the course of the equilibrium. A possible explanation of the observed change in q is suggested.

Introduction

The conductivity of air in the lower atmosphere is largely governed by the concentration of small ions, which have mobilities of the order 1 $cm^2 sec^{-1} V^{-1}$. The conductivity λ is equal to $(n_1 e k_1 + n_2 e k_2)$ where e is the charge per ion, usually taken to be equal to the electronic charge; k_1, k_2 are the mobilities of the positive and negative ions, respectively; and n_1, n_2 are their respective concentrations. The generation of small ions is attributed to the effects of cosmic radiation, radiation from naturally-occurring radioactive elements, and some industrial processes (Mühleisen, 1953). The rate of generation, q, has a value near 10 ion pairs $cm^{-3} sec^{-1}$ over land. Destruction of small ions occurs by collision with other small ions, condensation nuclei or droplets. The first effect is small compared with the other two in industrial areas. A measure of the second effect is the concentration of condensation nuclei in the range 4×10^{-7} to 2×10^{-5} cm radius, as recorded by an Aitken type counter. The usual symbol for this quantity is Z cm^{-3} and an average value for Oxford air is 20,000 cm^{-3}. The larger droplets are not usually present in large numbers and their effect is neglected. As the large ions formed by collision between small ions and condensation nuclei have very low mobilities (about 0·0003 $cm^2 sec^{-1} V^{-1}$) their contribution to the total conductivity is small except where they are present in large concentrations.

Experimental confirmation of the precise relationship between Z, n and q has been sought by many workers (e.g. Nolan, 1932; Sagalyn and Faucher, 1956; Torreson, 1939) and whilst qualitative agreement is found, results differ greatly in detail. From theoretical considerations Schweidler (1918) proposed the equation for the rate of growth of positive small ion concentration:

$$\frac{dn_1}{dt} = q - \alpha n_1 n_2 - \eta_{10} n_1 N_0 - \eta_{12} n_1 N_2 \qquad (1)$$

where N_0, N_2 are the numbers of uncharged and negatively charged condensation nuclei per cubic centimetre, respectively, and α, η_{10}, η_{12} are constants for a given

* Now at Marlborough College, Wiltshire.
† Now at the National Research Council of Canada.

VOL. 17 1959/6

type of ion and nucleus, but which vary with the radii of the particles concerned. Such constants are termed combination coefficients. α has a value in air of approximately $1{\cdot}6 \times 10^{-6}$ cm^{3}sec^{-1} at 20°C and 76 cm of mercury pressure. Values given for η_{10}, η_{12} show a wide scatter but they are of the order 10^{-6} cm^3 sec^{-1}. Determinations of the latter two constants depend on measurement of the ratios n_1/n_2 and N_1/N_0 or N_2/N_0 for which theoretical expressions involving the η's have been suggested. For a complete determination, however, it is necessary to make some assumption about the relative values of the η's. In what follows it will be assumed that:

$$\eta_{10} = \eta_{20} = \eta_0 \qquad \text{and} \qquad \eta_{12} = \eta_{21} = \eta$$

Various sets of observations of the equilibrium condition, as summarized by Nolan (1956) in a comprehensive survey of the present state of this subject, appear to fit a curve of the type

$$q = \alpha n^2 + \zeta n Z^x \tag{2}$$

where values between 0·2 and 1·0 have been reported for x.

Reliable data has probably been obtained by Holl and Mühleisen (1955), who have developed instruments for recording λ and Z continuously. Other observers have used Aitken counters or photoelectric counters of the type described by Nolan and Pollak (1946), together with a Gerdien type apparatus for conductivity measurement. Nolan (1956) concludes that the most likely value for x is 0·5 and gives several ways of accounting for this theoretically. It is not surprising that earlier experiments in which some of the quantities appearing in the equilibrium equation were measured should show a great scatter of results, since the difficulty of obtaining representative air samples and the inconvenience of the original Aitken counter provide sufficient reason. It is perhaps less expected that such a scatter should persist even with the modern techniques used by Holl and Mühleisen, whose results still show only a statistical relationship between Z and n rather than a clear verification of theory. Nolan (1950) has given the explanation that equilibrium is never reached in nature, so that attempts to determine the precise equilibrium relationship are of no use.

Indoor Measurements

Much work on the behaviour of nuclei stored in small gasometers has been done by the Nolans and although direct measurements of n_1, n_2 have not been made, because the sampling volume of the usual instruments is too large, these quantities have been deduced from N_1, N_2. In this way indirect checks on the equilibrium equation and constants have been made. Such experiments are only quasi-static since Z is continually decreasing by coagulation and fall-out of the nuclei: the average radius of the nuclei will also change and probably, as will be shown later, q is varying.

Two main reasons have been given for doubting the utility of further study of the equilibrium condition: first that equilibrium does not exist in nature, and hence outdoor measurements are not capable of giving more precise data than is already available; second, indoor conditions are very different from those outdoors,

hence any results obtained will not be applicable to natural conditions. The present writers are of the opinion that the value of the sealed-room type of experiment lies in the possibility of checking relatively simple equations and consequently the basic assumptions made in deriving them. For example, sizes and size-distributions of atmospheric nuclei have been measured by various methods, but it is not certain that every collision between a small ion and a nucleus results in combination. This means that the effective nuclear radius deduced from equilibrium measurements will probably not agree with those deduced from electron microscope measurement, from tuned collector experiments (Wilkening, 1952) or from diffusion experiments (Nolan and Doherty, 1950). A detailed knowledge of such disagreement may lead to a better understanding of the nature and behaviour of the fundamental particles of atmospheric electricity.

The content of the present paper is a report of an attempt at the complete determination of the form of the equilibrium equation which, although not completely successful, indicates that the method is capable of greater precision than outdoor experiments and points out some of the difficulties to be overcome in future attempts. The approach used was to make measurements of as many of the quantities appearing in the equilibrium equation as possible. Apparatus available was a conductivity apparatus, which could be used to record the concentration of small ions of either sign, a Nolan–Pollak type photoelectric counter for condensation nuclei, and standard instruments recording pressure, temperature and humidity. The room used, which had a volume of 55 m^3, was completely lined with sheet copper except for the window and contained a certain amount of electronic equipment. Although it was sealed with considerable care the internal barometric pressure tended to follow the outside trends. All experiments were conducted when the external pressure was decreasing: the diurnal temperature fluctuation was less than 0·5°F and the mean temperature decreased at the rate of 0·5°F per day. It is not considered that appreciable errors resulted from leakage of polluted air into the room. After the first two runs it was apparent that some further factor was varying with time and a simple apparatus was assembled to measure q. The q values obtained are of low accuracy but the qualitative variation found is consistent with the variations in n and Z.

VOL.
17
1959/6

Theory

The equation investigated was:

$$q = \alpha n^2 + \zeta n Z^x \tag{2}$$

The two most likely cases being $x = \frac{1}{2}$ and $x = 1$. In the latter case ζ becomes a constant β, which is thought to depend only upon the radius of the nuclei. Since it was not possible to make direct measurements of the mean radius of the nuclei it was proposed to use two fairly well established theoretical relationships.

(a) From previous work on the decay of nucleus concentration in small (300 l.) sealed containers the equation

$$\frac{dZ}{dt} = -\gamma Z^2 - \lambda^1 Z \tag{3}$$

where γ and λ^1 are constants, had been suggested (NOLAN, 1950). Recent measurements by SAGALYN and FAUCHER (1956) have shown that this equation applies approximately to the free atmosphere at various altitudes. The first term on the right-hand side represents loss by coagulation and γ may be calculated using SMOLOCHOWSKI's formula: $\gamma = 8\pi Dr$, where r is the radius of the particles considered and D is a diffusion coefficient. The second term represents loss by fall-out and diffusion to the boundaries of the container.

(b) Neglecting the loss of water by diffusion and sedimentation we may assume $Zr^3 =$ constant. Combining these relationships it is possible in principle to compute the value of r at any time. In the present work initial values only were determined since the uncertainties in q are greater than the corrections for variation of r which should be applied to equation (3).

By solving the differential equation

$$\frac{dn}{dt} = q - \alpha n^2 - \beta n Z$$

it is possible to obtain the time constant τ characterizing the change in n when q or Z are suddenly changed. Assuming $q = 10$ cm^3 sec^{-1}, $\beta = 3 \times 10^{-6}$ cm^3 sec^{-1} and $\alpha = 1{\cdot}6 \times 10^{-6}$ cm^3 sec^{-1}, which are values given by other workers, we have for

$Z = 10^5$ cm^{-3}	$\tau =$ 3 sec
10^4	30
10^3	300

So that, for the relatively smooth variations in Z and q with time observed in these experiments, the equilibrium condition should be closely followed.

DESCRIPTION OF APPARATUS

Measurement of n_1 *and* n_2

n_1 and n_2 were measured by collecting all ions of mobility greater than 0·2 cm^2 sec^{-1} V^{-1} in the air sampled by a condenser apparatus. Air was drawn at 6·1 l. sec^{-1} through a large parallel plate condenser which had a capacity of about 250 pF. One set of plates was maintained at $\pm 9{\cdot}3$ V with respect to the set of collector plates. All ions having mobility k greater than k_{crit} were collected and the current associated with them was measured. The critical mobility for this type of apparatus is given by

$$k_{\text{crit}} = \frac{\phi}{4\pi C V}$$

where ϕ is the air flow between the plates, C is the capacity between deflectors and collectors and V is the potential difference between them. If each ion carries unit electronic charge, the current flowing to the collectors is simply

$$i = ne\phi$$

This current was measured with a standard vibrating reed amplifier, having an

input resistance of $10^{11}\ \Omega$ and which operated an Elliott pen recorder. The measurement of n_1 and n_2 occupied about 3 min so that a relatively large sample of room air was used in each measurement. The small ions removed by the apparatus were rapidly replaced by the radioactivity in the room. Only a slight inaccuracy was caused by the unavoidable collection at the same time of a proportion k/k_{crit} of nuclei having mobilities less than $0{\cdot}2\ \text{cm}^2\ \text{sec}^{-1}\ \text{V}^{-1}$. This method of total collection was considered preferable to the more usual method whereby n_1 and n_2 are estimated from the measured value of air conductivity, using

$$\lambda_1 = n_1 e k_1$$

where some assumption concerning k_1 is necessary.

Measurement of Z

A photoelectric nucleus counter as described by Nolan and Pollak (1946) was used to measure Z. Three successive expansions were made for each reading and readings were taken at intervals of the order 2 hr. The sample of air passed into the counter through 1 ft of 0·5 in. inside diameter brass tubing and 1 ft of 0·5 in. rubber tubing. The brass tube projected into the room to prevent the sampling of air too close to the wall.

Measurement of q

Measurement of q is difficult and the method described here is not claimed to give anything more than an indication of the manner in which q changes.

A cylindrical ionization chamber was constructed from metal gauze. A central electrode of 0·28 cm diameter screwed rod was connected to the input of a vibrating reed amplifier. The outer cylinder was maintained at +48 V with respect to the rod. The current i to the central electrode is a measure of q:

$$i = qVe$$

where V is the chamber volume and e is the electronic charge. A slow current of air through the chamber ensured that the nucleus content was the same as that in the rest of the room. The number of ions carried into the chamber was insufficient for the conduction current to contribute appreciably to i. Three objections to this method are:

(a) The value of q given is a localized value at the site of the chamber and is affected by the radioactivity of the chamber itself.

(b) Since the ions are removed rapidly from the α-particle tracks there will be less recombination in these tracks than occurs in room air. This effect leads to a spuriously high q-value.

(c) The chamber walls intercept some α-particles which would normally contribute to the ionization inside the chamber volume: this will tend to give a low q-value.

Correction was made for the radioactivity of the chamber itself by determining the current i when the chamber was filled with nucleus free nitrogen and surrounded by lead blocks. The current i so found corresponds to the "zero q" line marked on the q-time graphs. An Elliott recorder connected to the output of the vibrating reed amplifier gave a continuous indication of the variation of q with time.

VOL. 17 1959/60

Accuracy

Probable errors in Z are considered to be:

for $Z > 3000$, $\pm 2\%$

$3000 > Z > 1500$, $\pm 3\%$

in addition to possible calibration and sampling errors.

The probable error in n is $\pm 5\%$ apart from sampling errors. No correction has been made for the presence of ions having mobilities less than $0{\cdot}2$ cm^2 sec^{-1} V^{-1}. All ions were assumed to carry a single electronic charge. No correction was made for the end effect of the condenser apparatus which was thought to be negligible.

Results

Five experiments were made, two during August and September 1955 and three in February 1956. During the last three runs a measurement of q was made in

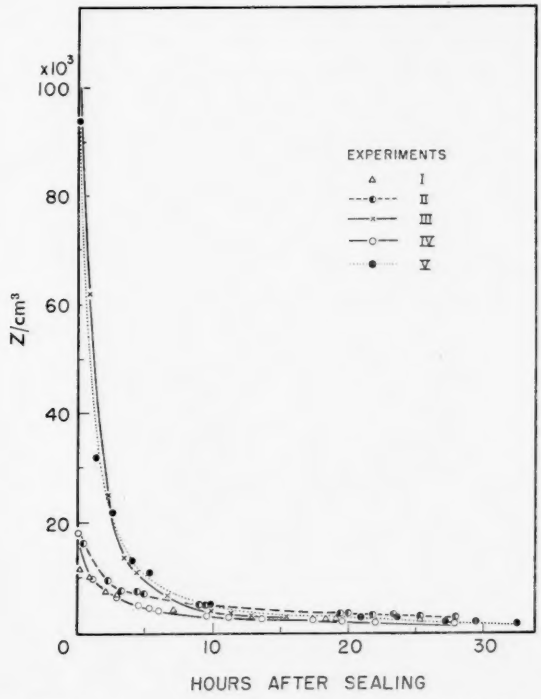


Fig. 1. The decay of condensation nucleus concentration with time for five samples of air inside a sealed room.

addition to measurements of n_1, n_2 and Z. The over-all picture of events is remarkably consistent and it is interesting to note that in experiment III artificial nuclei were used (because the external Z-count was unusually low at the time of starting the experiment). A small piece of cotton wool was soaked in alcohol and burned, shortly before the room was sealed.

Figs. 1 and 2 show the variation of Z and n_1 with time. Fig. 3 shows a plot of

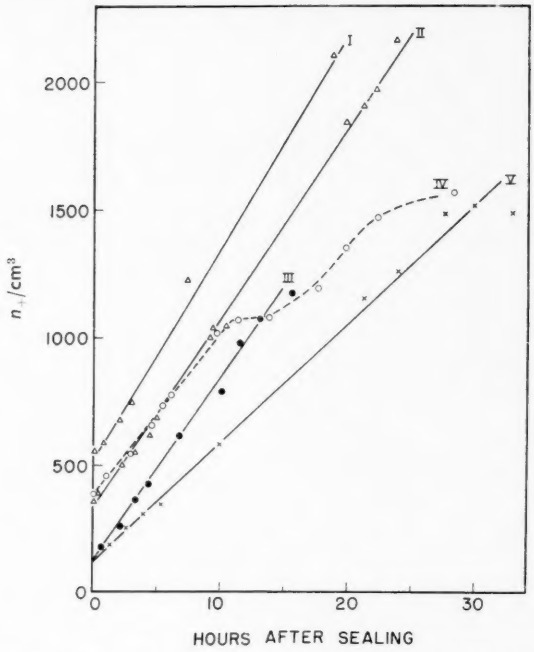


Fig. 2. The increase in small ion concentration with time inside a sealed room. The curves for n_- were similar.

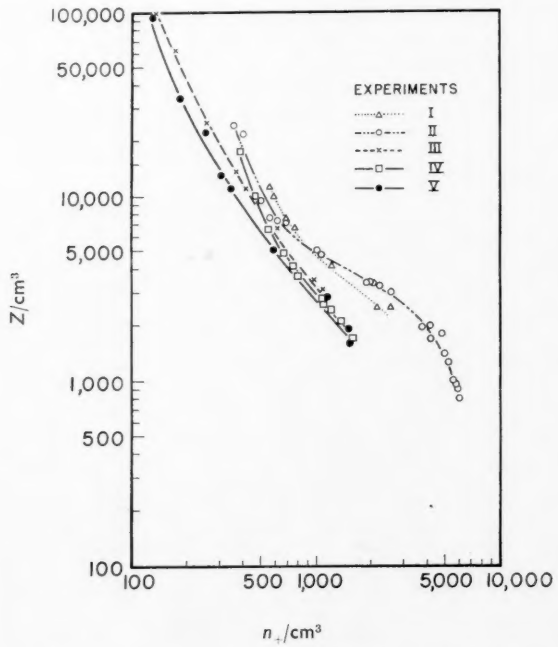


Fig. 3. Logarithmic plots of Z against n_+ for five samples of air.

VOL.
17
1959/6

n_1 against Z using logarithmic scales; the slope, for large values of Z, is a measure of x (equation 2 above). Fig. 4 shows the variation in q recorded during experiments III, IV and V; points shown are for experiment III only.

A value for the constant λ^1 (equation 3 above) was found as follows: by trial and error the value of λ^1/γ to give the best straight line for a plot of log

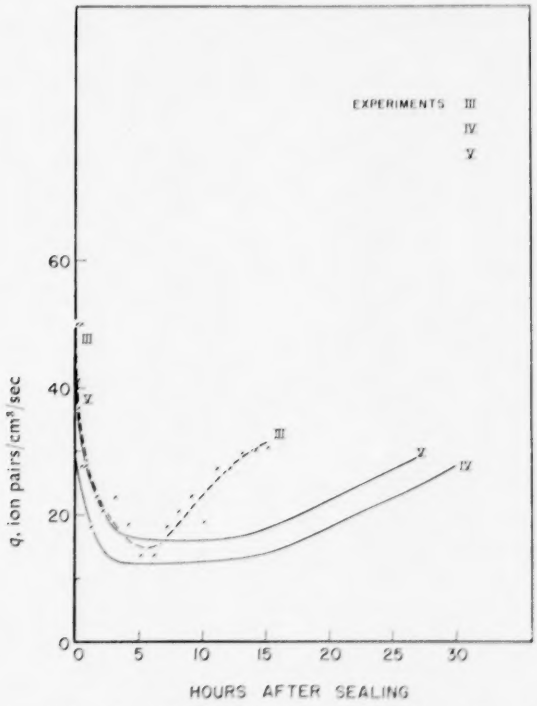


Fig. 4. The variation of q with time for three samples of stored air. The individual points are values averaged over several minutes and apply to experiment III only. A similar scatter was observed in each of the other two experiments.

$(Z + \lambda^1/\gamma)/Z$ against time was found to be 500, but this value is too small for the method to be sensitive and accuracy may not be better than ± 50 per cent. Values of λ^1 are not of special interest, since they depend on the dimensions and nature of the container.

Discussion

(a) *Decay of nucleus concentration with time*

The values of Z and t are found to fit the equation

$$\frac{dZ}{dt} = -\gamma Z^2 - \lambda^1 Z \qquad (3)$$

as given by Nolan, but the values computed for γ are higher than those reported by Kennedy (1916) and by Nolan (1941). Kennedy observed nuclei emitted

from a Bunsen burner flame and stored in a gasometer which was water-sealed. It is possible that such nuclei would be considerably larger than those occurring in the free atmosphere and that the value for γ would be correspondingly lower. Hogg (1934) has also reported a high value for γ for natural nuclei. Since the gasometer used by Nolan was oil-sealed there seems to be no reason for unusually rapid growth of the stored nuclei, and it must be concluded that the different γ values mean that the Dublin nuclei are larger than those in Canberra and Oxford. The equation suggested by Smith and Schilling (1954):

$$\frac{dZ}{dt} = -\psi Z^{3/2}$$

does not fit the results as well as the Nolan equation (3) above.

(b) *Variation of* n

From Fig. 2 the value of n_1 is seen to increase linearly with time for the first few hours of each experiment. The variation of q appears to preclude any clear cut deduction from this observation.

It is interesting to note that the value of n_1/n_2 did not increase steadily to a maximum at 1·4, as it did in the measurements made by Gunn and Phillips (1957): the lowest value of Z in the present work is considerably larger than those used by Gunn.

(c) *Change in* q

The value of q varies considerably with time, as shown in Fig. 4. Owing to uncertainties in the absolute value of q, it does not seem possible to use these results in a closer check on the equilibrium equation. A possible explanation of the shape of the curve is that two processes are involved. The initial fall in q may be attributed to a loss of radioactive elements from the air, perhaps associated with the initial rapid fall in Z. The evidence of Burke and Nolan (1950) that, except for very high concentrations of condensation nuclei, the radium-A content of the atmosphere is mainly associated with the intermediate ions, may be significant here. Subsequently a slow diffusion of radioactive gases from the walls of the room may be the dominating influence until an equilibrium concentration is approached, so that q increases and eventually becomes constant. A rough calculation shows that the rate of exhalation of radon from unit wall area would require to be some twenty times the value given for a soil surface by Chalmers (1957). This sort of variation could account for the observed change in q and for the later stages of experiment II shown in Fig. 3.

(d) *The check on the equilibrium equation*

The graph of log Z against log n_1 shows that the points for each experiment lie on a separate curve with little scatter. An average value for x of 1/2·2 may be obtained from the initial slopes, which compares with the value 1/2·3 as found by Holl and Mühleisen at Stuttgart and Basel. The curvature of the lines cannot be accounted for by the extra term (αn^2) which is small, but it may be related to the q variation mentioned in (c) above. Since q falls for the first 2–4 hr the initial slopes

VOL.
17
1959/6

will appear too steep, and as q increases the slopes will be too shallow. The values of x given are below the correct value on this view.

A plot of n^2 against $(nZ^{1/2})$ or (nZ) is not the line of nearly constant slope which might be anticipated from the form of equation (2) in the absence of a change in q. Indeed, even the sign of the slope is not constant: it was this observation from the results of the 1955 experiments which led to the attempt to observe the behaviour of q directly in the later experiments.

(e) *Agreement with free air results*

The close agreement between the value for x found here and that reported by Holl and Mühleisen is not altogether expected. The latter authors explain their value for x by assuming that the relationship between mean radius and number of nuclei (Junge, 1952)

$$\frac{dZ}{d(\log r)} = \frac{C}{r^3}$$

applies to surface air near an industrial area. It is by no means certain that such an assumption is justified, without experimental verification at the site used in the experiments. The observed fluctuations in Z from minute to minute could be caused by vagaries of wind bringing air from several sources of pollution, each source being characterized by its own radius spectrum and level of radioactivity. In a sealed room experiment, $Zr^3 =$ constant seems to be a fair assumption. However, the experiments of Sagalyn and Faucher have shown that, in relatively clear air far from industrial areas, coagulation of nuclei on a large scale is a major factor in determining the electrical structure of the lower troposphere. The difference in sampling technique between air-borne and ground experiments is, of course, considerable, but here is at least an indication that $Zr^3 =$ constant may be applied in the free atmosphere in some situations. Additional support for Holl and Mühleisen may be deduced from the similarity between experiment III using artificial nuclei and the rest of the experiments in the present series, indicating that the origin of the nuclei may not have too great an effect on their size distribution and hence on their atmospheric electrical behaviour.

OL.
17
9/60

Conclusions

(i) Study of the precise form of the equilibrium equation can lead to a more exact knowledge of the physical nature of atmospheric ions and nuclei.

(ii) Experiments on nuclei stored in large containers are capable of yielding more precise results than open air experiments.

(iii) In future work an accurate knowledge of the magnitude of q at all times is essential: the value of q is not constant in a sealed room.

(iv) The relationship between q and Z should be more closely examined. In the experiments described here the variation in q controls the equilibrium condition. It would be interesting to repeat the outdoor measurements of Holl and Mühleisen, employing an additional apparatus to record short-term fluctuations in q. It is possible that such fluctuations are responsible for the large scatter in their results rather than, as Nolan suggests, the absence of the equilibrium condition in nature.

(v) The present experiments yield the following results:
The equation for decay of nucleus concentration is

$$\frac{dZ}{dt} = -\gamma Z^2 - \lambda^1 Z$$

where γ has a value 6×10^{-9} cm^3 sec^{-1}, corresponding to a calculated mean radius of 6×10^{-7} cm for Oxford nuclei. The initial form of the equilibrium equation is

$$q = \alpha n^2 + \zeta n Z^{1/2.2}$$

Acknowledgements—We should like to express our gratitude for the use of the facilities of the Clarendon Laboratory and to Mr. Barklie of the Admiralty Research Laboratory, Teddington, for arranging the loan of some of the apparatus used.

The work was financed by an Admiralty research grant.

References

Burke T. and Nolan J. J.	1950	*Proc. R. Irish Acad.* **52,** 145.
Chalmers J. A.	1957	*Atmospheric Electricity* p. 80. Pergamon Press, London.
Gunn R. and Phillips B. B.	1957	*J. Meteorol.* **14,** 272.
Hogg A. R.	1934	*Gerlands Beitr. Geophys.* **41,** 32.
Holl W. and Mühleisen R.	1955	*Geofis. Pur. Appl.* **31,** 115.
Junge C.	1952	*Ber. dtsch. Wetterd.* **35,** 261.
Kennedy H.	1916	*Proc. R. Irish Acad.* **33,** 58.
Mühleisen R.	1953	*Z. Geophys.* **19,** 142.
Nolan P. J.	1956	*J. Atmosph. Terr. Phys.* **9,** 295.
Nolan P. J. and Doherty D. J.	1950	*Proc. R. Irish Acad.* **53,** 163.
Nolan P. J. and Pollak L. W.	1946	*Proc. R. Irish Acad.* **51,** 9
Nolan P. J.	1941	*Proc. R. Irish Acad.* **47,** 25.
Nolan P. J.	1932	*Proc. R. Irish Acad.* **41,** 61.
Nolan P. J. and Nolan J. J.	1932	*Proc. R. Irish Acad.* **41,** 111.
Sagalyn R. and Faucher G.	1956	*Quart. J. R. Met. Soc.* **82,** 428.
Schweidler E. von	1918	*S.B. Berl. Akad. Wiss.* **127,** 953.
Schweidler E. von	1918	*S.B. Berl. Akad. Wiss.* **128,** 947.
Smith L. G. and Schilling G. F.	1954	*J. Atmosph. Terr. Phys.* **4,** 314.
Torreson O. W.	1939	*Terr. Mag. Atmos. Elect.* **44,** 59.
Wilkening M. H.	1952	*Rev. Sci. Instrum.* **23,** 13.

VOL 17 1959/

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 145 to 149. Pergamon Press Ltd. Printed in Northern Ireland

Rapid fluctuations during magnetic disturbance

J. Lawrie
Dept. Scientific and Industrial Research, Magnetic Survey, Botanic Gardens, Christchurch C.I., New Zealand

(*Received* 16 *June* 1959)

Abstract—A numerically simple ratio is defined and used to examine space relationships of rapid geomagnetic fluctuations during disturbance.

Introduction

It was thought helpful towards the study of rapid fluctuations during magnetic disturbance to try to establish the degree of relationship between fluctuations simultaneously recorded at different points.

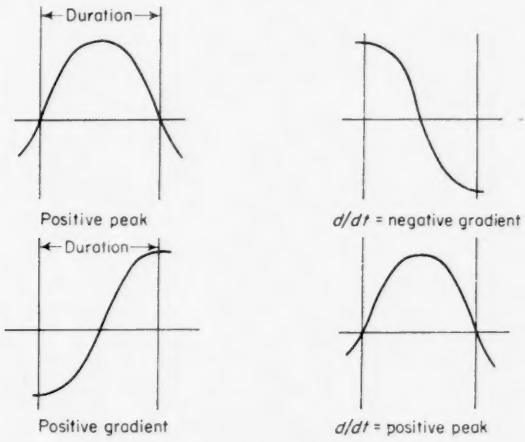


Fig. 1. The two basic types of fluctuation and their differentials.

The storm of 9–13 November 1956 was selected, making use of rapid rate records. Lovö was chosen as the basic station, and about 100 marked impulses in H there were picked, ranging from 30 sec to 5 min in duration. The impulses have a wide variety of recorded forms, and must include a number of *pt*, *pc* or *pg* type. In fact it is very difficult to distinguish random fluctuations from recognized types during marked disturbance, and initially no such classification was attempted.

Coincidence Ratio

Each fluctuation at Lovö was described by: (a) type—either peak or gradient; (b) duration; (c) time—at the centre of the fluctuation; (d) sign. (See Fig. 1.)

Inspecting one of the components at a particular station, a quantity ω_r was assigned for the rth incident as follows:

$\omega_r = +1$ or -1 : an incident of the same type, of the same or opposite sign, respectively, at the same time and of about the same duration;

$\omega_r = +\frac{1}{2}$ or $-\frac{1}{2}$: an incident of the same type, of the same or opposite sign, respectively, doubtful because of weak amplitude, or difference in duration, or time displacement up to a maximum of one-quarter of the duration;

$\omega_r = 0$: no similar incident for which the time displacement was less than one-quarter of the duration;

For the total N impulses

let $\Omega = (1/N)\sum\omega_r$

Then Ω_C is defined as the coincidence ratio for the component C at the station in question, relative to H at Lovö.

This ratio was evolved to avoid the particular difficulties involved in computing a correlation coefficient. It is considered to approximate to the correlation coefficient, differing mainly by being not necessarily reciprocal.

Method

Coincidence ratios relative to H at Lovö were calculated for both H and D at the stations listed in Table 1. The incidents were divided into four groups, each

Table 1. Stations used in the analysis

Station	Abbreviation	Geomagnetic co-ordinates	
		Latitude	Longitude
Lovö	Lo	+58·1°	105·8°E
Rude Skov	RS	+55·8	98·5
Wingst	Wn	+54·5	94·0
Witteveen	Wi	+54·2	91·0
Chambon-la-foret	CF	+50·4	83·9
Tamanrasset	Ta	+26·0	81·5
Hermanus	Hr	−33·3	80·5
Apia	Ap	−16·0	260·2

covering 6 hr of a Greenwich day, and the ratios calculated for each group. The results are expressed in Figs. 2 and 3, in which the stations are arranged in order of increasing distance from Lovö, but are not spaced according to any scale.

Coincidence ratios in H were calculated relative to dH/dt also at Lovö and these are shown in Table 2. They are for the whole day.

Results

As can be seen, the detail in H at Rude Skov and at Lovö is fairly similar. A substantial positive ratio persists as far as Witteveen and there is a positive

VOL. 17 1959/6

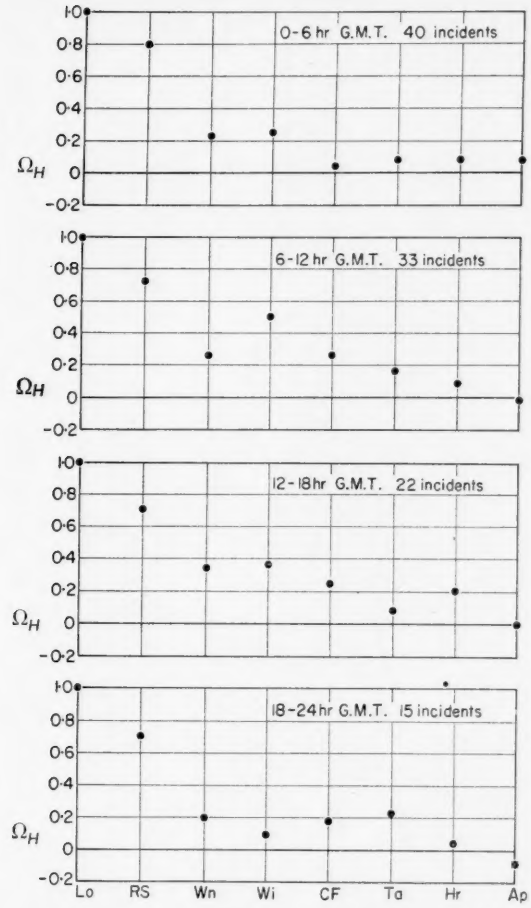


Fig. 2. Coincidence ratios in H relative to H at Lovö.

residue even at Hermanus. Although there are not many incidents in each group it is believed that these features are significant because of their persistence through all the groups and from station to station.

Lower ratios at Wingst result from the rate of change type of recording there, in which very short duration impulses mask longer duration ones.

The zero ratios at Apia are mainly because of the absence of any kind of coincidences there. For fluctuations of this order of rapidity, disturbed periods at Lovö are fairly calm at Apia and vice versa.

Now if simultaneous incidents at Lovö and another station were not related, there would be as many positive as negative values of ω_r, and Ω should be approximately zero.

Again, many of the impulses which have a positive H departure at Lovö must be associated with negative departures elsewhere because of the current systems involved.

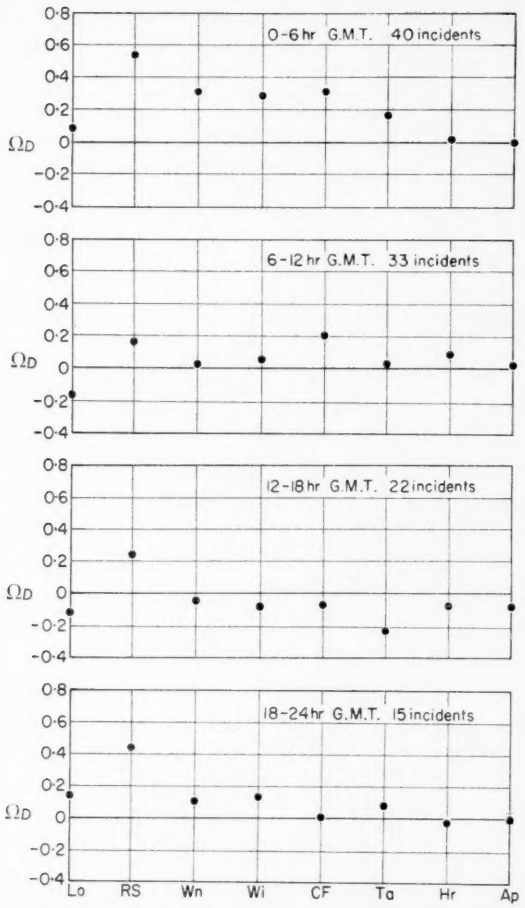


Fig. 3. Coincidence ratios in D relative to H at Lovö.

Table 2. Coincidence ratios in H relative to dH/dt at Lovö

RS	Wn	Wi	CF	Ta	Hr	Ap
+0·01	−0·02	0	−0·06	+0·05	+0·07	0

Thus it is surprising that Ω_H does retain a positive sign for such distances, and it must mean that a significant proportion of the impulses recorded simultaneously at least along the same meridian are functionally related.

It is to be expected that this tendency is produced by a certain class of the impulses considered. However, no clear dependence could be found on any of the following qualities: visual description, time of day, amplitude, duration and associated signs of departure in D and Z at Lovö.

VOL. 17 1959/6

Thus it appears that nothing short of examination of the whole spatial vector pattern will serve to establish a relationship between impulses recorded simultaneously at distant points.

The high value of Ω_H at Rude Skov shows that one can usually assume a relation between impulses in the group considered, simultaneously recorded over a distance less than 500 km.

The persistent feature of Ω_D is that $\Omega_D(RS) > \Omega_D(L_0)$. This probably indicates the curvature of the average current system in the vicinity, although it could perhaps be affected by earth induction effects as described recently by PARKINSON (1959).

All values of Ω_D have a maximum for 0–6 hr and a minimum for 12–18 hr. Because of the small sample this cannot be considered significant; but such a result is quite well to be expected from current systems having differing local patterns with time of day as with those corresponding to the longer duration impulses considered by VESTINE (1947, p. 269) and NAGATA and FUKUSHIMA (1954).

The negligible values of Ω relative to dH/dt show two interesting points.

Firstly, if there were a one way propagation involved so that events at Lovö tended to be associated with, say, later similar events at another point, this might show up in these ratios (see Fig. 1). The fact that there is no noticeable effect at Witteveen shows that any consistent phase difference there is less than 15 sec. Thus one way propagation along the line between Lovö and Witteveen must be on the average faster than 8×10^4 m/sec. Average speed of one way propagation along the meridian must be more than 3×10^4 m/sec.

Secondly, since no such effect is detectable, then functional relationships have been eliminated, and the residual values of Ω_H are determined by the random coincidences. This is then an empirical test of significance of the method, and the negligible values of Ω lend weight to the other results quoted above.

CONCLUSIONS

Although the network of stations considered is a limited one the following conclusions can be drawn for fluctuations of duration 30 sec to 5 min.

(1) A relationship can usually be assumed between simultaneous impulses recorded over a distance less than 500 km.

(2) A significant proportion of simultaneous impulses at all points at least near the same meridian are related; but the distinction of such impulses is not obvious and either whole vector patterns should be used to establish individual relationships, or statistical methods adopted.

(3) Over Europe any one way propagation along the geomagnetic meridian must be on the average faster than 3×10^4 m/sec.

Acknowledgements—The writer is indebted to A. L. CULLINGTON, Superintendent, and P. J. GILL of the Magnetic Survey for helpful discussions.

REFERENCES

NAGATA T. and FUKUSHIMA N. 1954 *Indian J. Met. Geophys.* **5**, 75.

PARKINSON W. D. 1959 *Geophys. J.* **2**, 1.

VESTINE E. H. *et al.* 1947 Publication 580 p. 269. Carnegie Institute of Washington.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 150 to 157. Pergamon Press Ltd. Printed in Northern Ireland

On the mode of propagation in the *E*-layer

W. R. PIGGOTT* AND J. BHATTACHARYYA†
D.S.I.R. Radio Research Station, Slough‡

(*Received* 17 *June* 1959)

Abstract—LÉPÈCHINSKY and DURANT's suggestion—that f_0E is incorrectly interpreted at stations where the dip exceeds about 60°—is critically tested using observed f_0E data. All the tests agree in showing that this hypothesis is wrong, the critical frequency being reduced correctly for stations with dips up to 89°. The common type of E_s trace extending approximately $f_H/2$ above f_0E, interpreted by LÉPÈCHINSKY as a magneto-ionic effect above f_xE, is shown to be due to sporadic E and not magneto-ionic effect. It follows that the collisional frequency near the maximum of the E-layer is of the order 10^4 sec^{-1} and not 10^6 sec^{-1} as required by LÉPÈCHINSKY.

INTRODUCTION

IN recent papers LÉPÈCHINSKY and DURANT (1955) and LÉPÈCHINSKY (1956) have challenged the correctness of applying the standard methods of deducing E-layer ionization density at stations where the magnetic dip exceeds about 60°. They suggest that the type of propagation near the point of reflection may change from quasi-transverse (Q.T.) to quasi-longitudinal (Q.L.) when the layer is low; for example, during the middle of the day in summer months. In addition LÉPÈCHINSKY (1956) claims that if the maximum of the layer is below the transition level between the Q.T. and Q.L. modes of propagation, a cusp should appear with a sporadic-E type trace extending to approximately $f_H/2$ above the cusp frequency, f_H being the gyro-frequency at E-layer heights.

In temperate latitudes $f_H/2$ is about 0·6 Mc/s and sporadic-E traces extending 0·5 Mc/s to 0·7 Mc/s above the cusp frequency are extremely common. For example, at Slough, England, during most months of the year, between 20 per cent and 40 per cent of the ionograms obtained near the middle of the day show traces obeying this criterion, and at Inverness in summer the proportion is often above 60 per cent.

It is therefore important to discover whether these traces are due to magneto-ionic causes, as suggested by LÉPÈCHINSKY, or to sporadic-E. Clearly if the propagation is really Q.L. in type, a fundamental change in our interpretation of the E-region phenomena will be essential.

We shall not, in this paper, discuss the relatively subtle points of magneto-ionic theory (LÉPÈCHINSKY, 1956, 1957) which determine whether LÉPÈCHINSKY's interpretation of the $h'f$ pattern is correct but will appeal directly to experiment. We may, however, note that this interpretation has been challenged by LANDMARK and LIED (1957) and does not appear to be consistent with the classical analyses of ECKERSLEY (1950) and RYDBECK (1950, 1951). We shall show that if LÉPÈCHINSKY's hypothesis is correct, the routine observations of E-layer critical

* Radio Research Station, Slough.
† Colombo Plan Scholar temporarily working at Radio Research Station, Slough.
‡ Official communication.

frequency, which are widely circulated and have been intensively studied in the past, should show abnormalities which are large compared with the errors of observation. These abnormalities have never been reported and we shall confirm, using limited random samples, that they do not exist. We shall also present evidence showing that the E_s traces involved are really due to E_s reflection and not to a magneto-ionic phenomenon.

2. General Theoretical Considerations

Lépèchinsky and Durant suggest that the collisional frequency in the E-layer in summer is of the order 10^6 per sec and that the lowest E critical frequency observed at temperate latitudes is, therefore, reflected in the longitudinal mode instead of the transverse mode, the second higher critical frequency occasionally seen being the transverse ordinary wave and not the extraordinary wave as is usually assumed.

Table 1

ν_c	10^7	3×10^6	10^6	3×10^5	10^5	3×10^4	10^4 sec^{-1}
θ	20°	43°	61·5°	73·5°	81°	85°	89°

The three possible critical frequencies are usually denoted by f_zE, f_0E and f_xE, respectively, the longitudinal mode being called the z-mode. We shall adopt this convention in this paper.

We may note that there are two ways in which reflections in both z and o-modes can be present together:

(a) by scattering of the wave in a direction parallel to the magnetic field (Ellis, 1953, 1956).

(b) by wave coupling between the z and o-modes at levels where the collisional frequency is near the critical value, ν_c, at which the propagation constants of the modes are equal (Eckersley, 1950 and Rydbeck, 1950, 1951).

The variation of amplitude with frequency produced by the two mechanisms and their incidence are fundamentally different though the z-mode critical frequencies are, of course, identical.

Apart from using values of f_zE produced by scattering mechanisms to establish the identity of f_0E and f_xE we shall be entirely concerned with the second mechanism which alone can give rise to the phenomena described by Lépèchinsky.

The transition level between Q.L. and Q.T. propagation occurs when the relation between the critical collisional frequency, ν_c, angle of dip, θ, and gyro-frequency, f_H, is given by

$$|\cos^2\theta/\sin\theta| = \nu_c/\pi f_H = a \tag{1}$$

or

$$\sin\theta = |\tfrac{1}{2}\{-a \pm \sqrt{(a^2+4)}\}| \tag{2}$$

Thus for $f_H = 1{\cdot}23$ Mc/s we find the corresponding values of θ and ν_c given in Table 1.

This, of course, agrees with Lépèchinsky's (1955) value of $\nu_c \doteq 10^6$ sec^{-1} at Poitiers, France, where the dip is 63°.

Clearly the point at issue is whether the values of the collisional frequency near the maximum of the E-layer are above or below these values.

The relations between the ordinary wave critical frequency for quasi-transverse propagation, f_0, the ordinary wave critical frequency for quasi-longitudinal propagation, f_z, and the extraordinary wave critical frequency, f_x, applicable for frequencies above the gyro-frequency are

$$\left.\begin{aligned} f_0^2 &= f_x^2 - f_x f_H \\ \text{and} \quad f_0^2 &= f_z^2 + f_z f_H \end{aligned}\right\} \tag{3}$$

giving
$$f_x - f_0 = f_x f_H/(f_x + f_0) \doteq f_H/2 \tag{4}$$

and
$$f_x - f_z = f_H \tag{5}$$

Thus the frequency separation between corresponding parts of the ordinary and extraordinary wave $h'f$ curves provides a definite distinction between Q.T. and Q.L. propagation. This is examined in Sections 3 and 4 below.

Near noon in summer months the variation of solar zenith angle χ with latitude is small so that the height of the E-layer is unlikely to change significantly with latitude. The level at which $\nu = \nu_c$ will, however, rise by 3·5 scale heights (about 25 km) when ν changes from 10^6 sec^{-1} to 3×10^4 sec^{-1} (dip 61·5° to 85°). Thus if the value of ν near the maximum of the E-layer, ν_m, is about 10^6 sec^{-1} in temperate latitudes, the E-layer at high latitudes will be so far below the level where $\nu = \nu_c$ that no transverse reflection will be possible and the separation of components will be f_H. If, however, ν_m is of the order 10^4 sec^{-1}, giving three components at high latitudes, there is no doubt that the reflection mode must be transverse at lower latitudes. The experimental evidence is considered in Section 3 below.

If we choose a value for ν_m intermediate between these extremes the mode of reflection will vary systematically with dip and with the height of the layer. As we shall see (Section 5) unless the mode remains constant throughout the day, discontinuities will appear in the diurnal plots of f_0E as a function of χ and unless it remains constant at noon throughout the year similar discontinuities will appear in the corresponding seasonal variations. None of these have ever been detected. Similarly if the mode changes at a given value of dip, critical frequencies measured at stations where the dip is greater than the critical value will be $f_H/2$ different from those taken where the dip is less than this value. These discrepancies, which are about 0·5–0·8 Mc/s, are enormously greater than the anomalies discussed in the literature (Allen, 1946, 1948; Harnischmacher, 1950; Appleton *et al.*, 1955; Beynon and Brown, 1956).

Turning to the E_s traces which obey Lépèchinsky's criteria we may note that if they are magneto-ionic in nature they will appear only at the z-mode critical frequency. If, however, the lower E critical frequency is f_zE and the trace is due to E_s the separation f_xE_s–f_zE_s will be f_H (equation 5). If, therefore, pairs of

E_s-traces are found which are separated by about $f_H/2$, Section 4, the identifications of f_xE_s, f_0E_s, f_xE, f_0E are certain. We shall also show that the variations of f_0E are so small that, unless a discontinuity in the diurnal variation of order $f_H/2$ can be established, the continuity of mode from noon to twilight cannot be challenged. Thus the twilight ionograms, which, at all stations, often show the complete 0 and x-traces for both E- and F-layers, firmly establish the normal identifications of the critical frequencies as f_0E and f_xE. While, at any station, it is possible to adopt a value of ν which would at least allow some of the data to be consistent with LÉPÈCHINSKY's hypothesis, this process will necessarily cause discontinuities to appear in the data of other stations which are not, in fact, found (Section 6).

3. Observations at High Latitudes

If we examine ionograms taken at high latitudes (WRIGHT *et al.*, 1957; WRIGHT and KNECHT, 1957) we find many examples showing all three components. In general, the z-component is weaker than the 0-component near their critical frequencies implying that $\nu_m < \nu_e$. The additional trace between f_zE and f_0E is, however, usually an F-reflection and not of the E_s type described by LÉPÈCHINSKY (RYDBECK, 1950, 1951; SCOTT, 1951; LANDMARK and LIED, 1957).

Thus, turning to SCOTT's (1951) paper on the gyro-frequency in the E-layer at very high latitudes, we note that his "typical example" of a very high latitude $h'f$ curve shows the E_0 trace stronger than the E_z trace near f_0E and f_zE, respectively, though weaker at lower heights and, in his text, he explains that the number of records showing f_0E, f_zE and f_xE are restricted owing to f_xE being missing in many cases—as forecast by LÉPÈCHINSKY (1955). Now both Baker Lake (dip 86·5°) and Resolute Bay (dip $\div$ 89°) should show solid longitudinal propagation with no 0-mode reflection if ν_m exceeds 10^5 sec^{-1}. Nevertheless SCOTT (1951) quoted 110 cases showing all three E-components in four months data from the former station, twenty-six of which occurred in June 1949. His figure 6 also shows forty-one cases with f_0E and f_zE readable together at the latter station during the three noon hours 1100–1300 of the summer months May to July 1949 and 1950. Thus Q.T. propagation has not been completely replaced by Q.L. even at these very high latitudes. We may conclude that ν_m is likely to be of order 10^4 sec^{-1} and that quasi-transverse propagation is possible for dip angles up to about 89°.

4. Comparison of Ordinary and Extraordinary Wave Critical Frequencies of E and E_s

Provided that the sensitivity of the ionosonde is sufficient and that it is suitably adjusted, the mode of propagation can be deduced by studying the frequency relations between corresponding parts of the ordinary and extraordinary wave traces. The adjustment of the ionosonde is usually rather critical; too great sensitivity gives a complex pattern of weak reflections in which the extraordinary wave trace cannot be identified with certainty, while too little sensitivity results in the weaker extraordinary wave traces being absent. Thus it is fairly easy to fail to detect both components regularly.

4.1. *Slough observations*

Ionograms obtained between 0800 GMT and 1600 GMT at Slough, dip 67°, for the two months June, July 1954, were examined for the presence of extraordinary wave traces reflected at or below the maximum of the normal E-layer.

Out of the 400 ionograms studied, 215 showed E_s traces to high frequencies due to dense E_s which completely prevented any separation of the two components. From the remaining 185 records, 150 showed both ordinary and extraordinary wave traces with the E-layer cusp frequencies measurable to $\pm 0{\cdot}1$ Mc/s or better. All but two of these gave separations between 0·6 Mc/s and 0·8 Mc/s, the remainder giving 0·85 Mc/s and 0·9 Mc/s. In addition, sixty-eight records gave usable doubled E_s traces. Again the differences found lay between 0·55 and 0·85 Mc/s. Three-quarters of these traces were of the type mentioned in the introduction, f_0E_s–f_0E being between 0·5 Mc/s and 0·8 Mc/s, with the majority giving f_0E_s–$f_0E \doteqdot 0{\cdot}6$ Mc/s. Since the E_s part of the patterns was visible on both ordinary and extraordinary wave traces there can be no doubt that they were caused by E_s phenomena and not magneto-ionic causes. We may conclude that the identifications of f_0E and f_xE have been made correctly.

4.2. *Inverness observations*

A brief examination of ionograms obtained at Inverness (dip 71°N, $\nu_c = 3{\cdot}3 \times 10^5$ sec^{-1}) during June 1955 showed that the recorder at this station was operating below the optimum sensitivity for observing the extraordinary wave reflection from the E-layer. Nevertheless, out of 204 hourly records obtained between 0800 and 1600 GMT, forty-seven showed measurable E-layer critical frequencies for both components. These gave separations averaging about 0·7–0·8 Mc/s. In addition, 100 records showed accurately measurable separations between the ordinary and extraordinary wave limit frequencies of sporadic-E. Thus once again the E_s traces themselves gave evidence that the 0·6–0·8 Mc/s extension of the trace was due to E_s and not magneto-ionic phenomena.

It is interesting to note that six of these records showed, in addition to the Q.T. ordinary and extraordinary wave traces, a very weak Q.L. (z-component) trace. This was much weaker than the corresponding part of the extraordinary wave trace suggesting that it was produced by scattering (Ellis, 1953, 1956), ν_m being much less than 3×10^5 sec^{-1}. However, if it was due to wave coupling, ν_m is likely to be less than 10^5 sec^{-1}. On three of these records F-layer Q.L. traces could be observed right down to f_zE, as seen at very high latitude stations. On these occasions there is no doubt that scattering at the angle of dip was more important than ν_c in producing the weak mode of reflection. The simultaneous presence of all three critical frequencies, of course, establishes the identity of the normal pair beyond any possible doubt.

5. Variability of the E-layer Critical Frequency

As is well known, the variability of the E-layer critical frequency is very small, the distribution curve normally being of the Gaussian type with standard deviations between 0·1 Mc/s and about 0·2 Mc/s. Taking, as an example, June

VOL.
17
1959/6

and July 1955, published observations between 1000 and 1400 LMT at Huancayo (dip 3°), Slough (dip 67°) and Churchill (dip 84°) give standard deviations about the monthly mean values of 0·11 Mc/s, 0·12 Mc/s and 0·23 Mc/s, respectively. Values for high latitude stations tend to be higher than those for low owing to the presence of short lived anomalies—most of which increase the apparent critical frequency, not decrease it as would z-reflections. The discontinuous nature of a change in mode would, therefore, be very striking and could scarcely be missed in even a cursory examination of the tables of data. Nevertheless a search

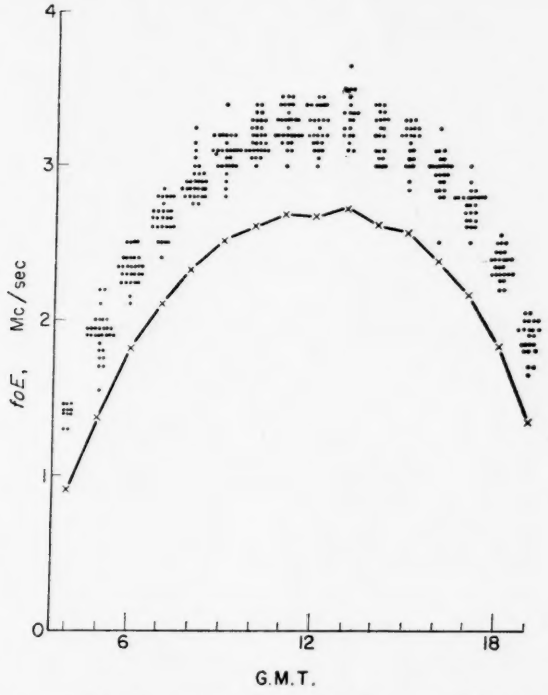


Fig. 1. f_0E at Slough, June 1955. Note: If $f_0E = 3{\cdot}0$ Mc/s, $f_xE = 3{\cdot}68$ Mc/s, $f_zE = 2{\cdot}44$ Mc/s,
×—×—× f_xE calculated from mean f_0E.

through tables of data from twenty-two stations has failed to disclose any effect of this type. Similar arguments show that it is only in exceptional circumstances that a change in mode would fail to cause a very distinctive discontinuity in the variation of monthly median or mean values of f_0E with position, season or time of day. No sign of these effects can be seen in ALLEN's (1946, 1948) graphs of f_0E as a function of sec χ. A typical scatter diagram for f_0E taken at Slough in June 1955 is shown in Fig. 1. It is obvious that one mode was present throughout the period. Since records taken when f_0E is near 2·0 Mc/s nearly always show very clear o and x-traces for both E- and F-layers there is no doubt that these values are interpreted correctly. We again conclude that the propagation was always of the transverse type.

6. Variation of f_0E with Latitude

Near noon in midsummer the zenith angle, χ, of the sun is almost exactly the same at Slough 51·5°N and Huancayo 12°S, $\chi = 35{\cdot}5°$ and $\chi = 37{\cdot}5°$, respectively. Thus the ionization density of the E-layer should be the same at these stations and the critical frequencies will be the same if both are reflected in the same mode. However, at Huancayo the dip is only 3° and there is no doubt that the propagation must be transverse. The mean E-layer critical frequencies observed at the two stations between 1000 and 1400 LMT during June and July 1955 are summarized in Table 2.

Table 2

Time	1000	1100	1200	1300	1400	Period
Huancayo	3·14	3·30	3·35	3·36	3·24	
Slough	3·19	3·26	3·25	3·32	3·18	June
Difference	0·05	0·04	0·10	0·04	0·06	1955
Huancayo	3·22	3·38	3·44	3·41	3·32	
Slough	3·19	3·23	3·29	3·31	3·24	July
Difference	0·03	0·15	0·15	0·10	0·08	1955

The standard deviation of the means for each hour are all close to 0·02 Mc/s. No differences of the order 0·6 Mc/s appear and we may conclude that at both stations reflections are obtained in the same, transverse, mode. Since the observed E critical reflections at Poitiers and Slough agree with roughly the same accuracy we may conclude that Lépèchinsky's observations also necessarily agree with Huancayo.

The variation of f_0E with latitude at noon in midsummer is relatively small, about 0·3 Mc/s, whereas the separation f_0E–f_zE varies from 0·5 Mc/s to 0·8 Mc/s.

Turning to the tables of data circulated by the Canadian Defence Research Board, we find that the tables of f_0E at Resolute Bay (dip 89°) for June and July 1954 include 475 numerical values between 0800 and 1600 LMT out of a possible 549, show very little scatter—certainly much less than $f_H/4$, and give values which are consistent with those obtained at Slough provided it is assumed that both stations are observing the Q.T. ordinary wave component f_0E. Similar results are obtained at Baker Lake (dip 86·5°). Since all three components are seen at these stations, this gives an independent check that Slough is measuring f_0E and not f_zE. Correcting the small differences due to changing zenith angle, χ by assuming that $(f_0E)\sec^{\frac{1}{2}}\chi$ is constant, we find, using over 700 observations in all, that corresponding mean critical frequencies at Huancayo, Slough and Churchill would be 3·51 Mc/s, 3·38 Mc/s and 3·55 Mc/s respectively. Again there is no evidence of any anomaly of the order $f_H/2$.

7. Conclusions

We conclude that the measured values of f_0E prove beyond reasonable doubt that the mode of propagation near the maximum of the E-layer is quasi-transverse

VOL. 17 1959/6

for all values of dip from 0° to 89° approximately; that the critical frequencies have, therefore, been correctly deduced throughout the world; that the collisional frequency near the maximum of E cannot greatly exceed 10^4 sec^{-1}, and that the common type of E_s extending about $f_H/2$ above f_0E is due to an E_s phenomena and not to magneto-ionic effects.

Although we have shown that the change in mode when $\nu = \nu_c$ does not affect f_0E, it is obvious that the changeover *must* occur either in the lower part of this layer or in the upper part of the D-region. The discussion of the effects produced, which are sometimes very evident, is, however, outside the scope of this paper.

Acknowledgement—The work described above was carried out as part of the programme of the Radio Research Board and the paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

References

Allen C. W.	1946	*Terr. Magn. Atmos. Elect.* **51,** 1.
Allen C. W.	1948	*Terr. Magn. Atmos. Elect.* **53,** 433.
Appleton E. V. Lyon A. J. and Turnbull A. G.	1955	*Nature, Lond.* **176,** 897.
Beynon W. J. G. and Brown G. M.	1956	*Nature, Lond.* **177,** 583.
Briggs B. H.	1951	*J. Atmosph. Terr. Phys.* **1,** 345.
Eckersley T. L.	1950	*Proc. Phys. Soc.* B **63,** 49.
Ellis G. E.	1953	*J. Atmosph. Terr. Phys.* **3,** 263.
Ellis G. E.	1956	*J. Atmosph. Terr. Phys.* **8,** 43.
Landmark B. and Lied F.	1957	*J. Atmosph. Terr. Phys.* **10,** 114.
Lépèchinsky D. and Durant J.	1955	*C. R. Acad. Sci., Paris* **240,** 333.
Lépèchinsky D.	1956	*J. Atmosph. Terr. Phys.* **8,** 297.
Lépèchinsky D.	1957	*J. Atmosph. Terr. Phys.* **10,** 247.
Rydbeck O. E. H.	1950	*J. Appl. Phys.* **21,** 1205.
Rydbeck O. E. H.	1951	*Chalmers. Tek. Högsk. Handl.* **101.**
Scott J. C. W.	1951	*J. Geophys. Res.* **56,** 1.
Wright J. W. Knecht R. W. and Davies K.	1957	*Annals of the I.G.Y.* Vol. III, p. 1. Pergamon Press, London.
Wright J. W. and Knecht R. W.	1957	*I.G.Y. Atlas of Ionograms.* C.R.P.L., Boulder, Colorado.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 158 to 169. Pergamon Press Ltd. Printed in Northern Ireland

The propagation of electromagnetic waves in ionized gases (with special reference to "whistlers")—I *

Introductory theory

F. H. NORTHOVER
Department of Mathematics, Carleton University, Ottawa, Canada

(*Received* 1 *June* 1959)

Abstract—Recent studies lend support to the theory that whistling atmospherics are caused by lightning flashes, the electromagnetic energy radiated by these being guided along discrete columnar ionic irregularities which follow approximately the lines of force of the earth's magnetic field. In Part I the theoretical problems arising are set forth and a general wave theory developed which is first applied to the problem of propagation through homogeneous compound streaming media. In Part II the simplest case of "standard type" propagation along stationary columns is carefully examined, both for columns with a central ionic surplus and for columns with a central ionic deficiency. Although both types of column, when sufficiently well developed, can guide electromagnetic energy it appears that the former type is a much more likely mechanism for the whistler propagation than the latter. It is hoped in subsequent papers to deal with the propagation of general type disturbances along stationary columns and also through columns where appreciable axial streaming is taking place.

VOL.
17
1959/6

1. INTRODUCTION

IT is now fairly well established† that whistling atmospherics are caused by lightning flashes, the electromagnetic energy radiated by these being guided fairly well along the geomagnetic field lines through the ionosphere.

This guidance alone does not seem sufficient to explain the high intensities sometimes observed in multiple or very long distance single paths but a waveguide effect might exist if the ionization were suitably bunched in columns along the field lines, and such an effect might explain the observed intensities. Further, this bunching could provide discrete preferred paths of propagation, each with its own travel time, which would account for the discrete dispersion traces frequently observed in the frequency against time of transit records of whistling atmospherics.

There is reason to believe that such columns might exist because electrons in a constant magnetic field have a natural tendency to execute helical paths whose axes are parallel to the magnetic field. Progression parallel to the field is at a constant rate and the number of circuits per second about the helix axis is equal to the gyro-frequency. The suggestion is, that in the upper reaches of the ionosphere, where collisions are relatively infrequent, this tendency causes inhomogeneities of charged particles to diffuse (in approximately helical paths) along some of the lines of force of the earth's magnetic field, ultimately forming more or less complete columns of ionization. It seems, therefore, worthwhile to investigate the propagation of electromagnetic waves in the presence of such columns.

* Contribution from the Radio Physics Laboratory of the Defence Research Telecommunications Establishment, Defence Research Board, Ottawa, Canada, under PCC D48-95-11-01, and from the Department of Mathematics, Carleton University, Ottawa, Canada.

† See "references".

In the applications to columns of the general analysis which will be developed in these papers, the curvature of the axis of the column will be neglected: we shall suppose that the column extends to infinity in both directions and shall neglect possible variations of ionic density in axial directions. As far as our wave treatment is concerned, this is equivalent to supposing that the axial curvature is very small compared with the reciprocal of the wavelength in the medium and that the percentage variations of ionic density along the axis are extremely small over many wavelengths.

It seems likely that the type of column to be considered consists of a compound cylindrically distributed ionic distribution consisting of a neutral ionized gas streaming through a neutral background of stationary ionized gas. If N is the ionic density of the moving gas and $\bar{N}$ this quantity for the stationary gas, U the stream velocity of the moving gas and C the velocity of light, then the case of practical interest re. whistlers is that of $N/\bar{N}$ small, U/C small. If U/C is sufficiently small we have effectively a stationary columnar distribution, and, in view of the extreme difficulty of the whole problem we shall make the first application of the general theory as regards columns to this case.

"Plane wave incident" (standard) type disturbances

Considering, for the moment, the modes of propagation down a stationary homogeneous cylindrical column (N constant through the column, N equal to a different constant outside), the disturbance within the column as the radius of cross-section tends to infinity must approximate to those possible in a homogeneous medium of infinite extent. The simplest of these latter types of disturbance is the plane wave-mode with direction of travel parallel to the original column's axis. The disturbance, propagated down the actual column, which could tend to this simple type were the column radius to tend to infinity will be called the "plane wave incident" or "standard" type disturbance. It will be taken as the standard type of disturbance in the present theory.

The "standard" disturbance propagated down a non-homogeneous axially symmetric ionized column would be defined as the one which would reduce to the above basic axially travelling plane wave, when $dN/dr \to 0$ for every r, r being the distance from the axis of a general point in the distribution.

In the more complicated case of the compound distribution, it appears that plane wave propagation is in general possible when both constituent distributions are homogeneous. "Standard" type disturbances can therefore be defined in a similar fashion for them. The theory would therefore seem to fall naturally into the following divisions:

Propagation down a simple columnar distribution. Standard and non-standard type propagation in non-homogeneous and homogeneous columns.

Propagation down a compound columnar distribution. Standard and non-standard propagation down non-homogeneous and homogeneous columns.

After general equations governing wave propagation of electromagnetic disturbances through ionized gases have been developed, the writer will endeavour to deal with some at least of these problems in a series of papers.

The original magneto-ionic theory as developed by APPLETON and HARTREE is

not sufficiently general for our requirements. We shall therefore employ magneto-hydrodynamical methods. Since the simple type of distribution is a special case ($U = 0$) of the compound distribution, we shall begin by developing the general equations for the latter although the actual application will be first to the simple-type stationary distributions. This will obviate the necessity of developing a general type theory twice over.

2. General Theory of Waves in Conducting Gases

The instantaneous current vector $\mathscr{I}$ at any point due to the passage of the wave consists of four parts, namely, the contributions due to displacements of the positive and negative ions in the steadily streaming gas and the contributions due to displacements of the two kinds of ions in the "background" gas through which the former gas is flowing. The value of $\mathscr{I}$ thus calculated can then be used in MAXWELL'S equations of the field to obtain a wave-equation for the electric vector $\mathbf{E}$.

2.1. *Detailed consideration of the moving gas*

Let ρ be the gas density, $\mathbf{T}\,d\tau$ the resultant force upon a volume element $d\tau$, $\mathbf{W}$ the velocity of this element and Π the gas pressure. Then the hydrodynamical equations are*

$$\frac{D\mathbf{W}}{Dt} = \mathbf{T} - \frac{1}{\rho}\,\mathrm{grad}\,\Pi \tag{1}$$

where D/Dt denotes "differentiation following the fluid", i.e.

$$D/Dt = \partial/\partial t + \mathbf{W}\,.\,\mathrm{grad}$$

and

$$\mathrm{div}(\rho\mathbf{W}) = -\partial\rho/\partial t \tag{2}$$

Since we are supposing that the gas, although ionized, is electrically neutral, a volume element $d\tau$ (taken to contain a large number $2N\,d\tau$ of particles) contains $N\,d\tau$ electrons and $N\,d\tau$ protons.

The vibration of the electrons and protons is caused by the passing electromagnetic wave. The mass of a proton is considerably greater than that of an electron and so, for frequencies considerably greater than the proton gyro-frequency and not too near the proton plasma-frequency the wave is carried almost entirely by the electrons, the positive particles remaining almost stationary in comparison. We shall therefore concentrate mainly upon the propagation in the electron gas although for completeness the equations of the propagation in the complete medium will be written down. This would be of use in discussion of the frequencies excepted above.

We proceed to the derivation of the equations of the theory. The force per unit mass acting upon an element $d\tau$ either of electron or proton gas is compounded as follows:

* We assume that the gases are sufficiently rarified for the effects of viscosity to be unimportant, although, of course, the analysis could be carried through with the viscosity term present.

2.1.1. *The body force per unit mass,* $\mathbf{F}$ *say, due to non-electromagnetic causes.* In terrestrial problems this may be taken as the local value of "$\mathbf{g}$". This of course neglects the effect of the internal gravitational attraction of the gas on itself which may become important in cosmogonical problems but is not important here*.

2.1.2. *Collisional effects.* As is now well known, this can be represented by a friction type force on any given particle proportional to its velocity through the gas. If m is the mass of a negative ion, ν the collision frequency, $\mathbf{V}_i$ its velocity and $\mathbf{U}$ the velocity of the background material, then this effect can be allowed for by a force $-m\nu(\mathbf{V}_i - \mathbf{U})$. Summing over the particles in the volume element, the contribution of this type of force to the element $d\tau$ of electron gas is $-m\nu(\mathbf{V} - \mathbf{U})N\,d\tau$, where $\mathbf{V}$ is the average velocity of all the particles in the volume element. From considerations of momentum it is easy to see that $\mathbf{V}$ is also the effective (hydrodynamical) velocity of the gas element $d\tau$.

Similarly, denoting by "$\wedge$" quantities pertaining to the proton gas, the force on a volume element $d\tau$ of it due to this cause will be

$$-\hat{m}\hat{\nu}\,(\hat{\mathbf{V}} - \mathbf{U})N\,d\tau$$

We can remark here the connexion between the hydrodynamical velocity $\mathbf{W}$ of the element $d\tau$ of the complete (neutral) gas and the corresponding velocities $\mathbf{V}$, $\hat{\mathbf{V}}$ of the constituent negative and positively charged gases. From consideration of the total momentum of the element $d\tau$ of the neutral gas, we have

$$m\mathbf{V} + \hat{m}\hat{\mathbf{V}} = (m + \hat{m})\mathbf{W} \tag{3}$$

2.1.3. *Force produced on the element by the magnetic field.* Considering first the negatively charged gas only, the force exerted by the magnetic field $\mathbf{B}$ upon a constituent particle is (neglecting relativistic effects) $-(e/C)\,\mathbf{V}_i \wedge \mathbf{B}_i$, $\mathbf{B}_i$ being the magnetic induction vector at the ith particle. Summing this over the particles in the element $d\tau$, the total force produced on the element $d\tau$ of electron gas through this cause, is $-(e/C)N\,d\tau(\mathbf{V} \wedge \mathbf{B})$, $\mathbf{B}$ denoting the magnetic field at the element.

The corresponding force for the positively charged gas will, of course, be

$$+(e/C)N\,d\tau(\mathbf{V} \wedge \mathbf{B})$$

2.1.4. *Force due to the electric field* $\mathbf{E}$. This is obviously $-(Ne\,d\tau)\mathbf{E}$ for the element of electron gas and $+(Ne\,d\tau)\mathbf{E}$ for the proton gas.

The hydrodynamical equation of motion for the electron gas may now be written down:

$$\frac{D\mathbf{V}}{Dt} = -\frac{e}{m}\mathbf{E} - \nu(\mathbf{V} - \mathbf{U}) - \frac{e}{mC}\mathbf{V} \wedge \mathbf{B} - \frac{1}{mN}\operatorname{grad} p + \mathbf{F} \tag{4}$$

where p here is the electron pressure (and so the force it exerts on $d\tau$ is $-d\tau$ grad p). Likewise the corresponding equation for the proton gas is

$$\frac{D\hat{\mathbf{V}}}{Dt} = \frac{e}{\hat{m}}\mathbf{E} - \hat{\nu}(\hat{\mathbf{V}} - \mathbf{U}) + \frac{e}{\hat{m}C}\hat{\mathbf{V}} \wedge \mathbf{B} - \frac{1}{\hat{m}N}\operatorname{grad} \hat{p} + \mathbf{F} \tag{5}$$

where $\hat{p}$ is the proton gas pressure.

* A note on the calculation of this effect will be given later.

Since the electron gas density is mN the hydrodynamical equation of continuity is

$$\text{div } (N\mathbf{V}) = -\partial N/\partial t \tag{6}$$

But the current vector $\mathbf{I}$ of the electron gas is $-eN\mathbf{V}$ and the electric charge in the volume $d\tau$ for this gas is $-eN\,d\tau$. Hence the electrical equation of continuity is the same as the hydrodynamical one and we shall therefore refer to (6) without ambiguity as "the equation of continuity" (for the negatively charged gas).

Similarly the equation of continuity of the proton gas will be

$$\text{div } (N\hat{\mathbf{V}}) = -\partial N/\partial t \tag{7}$$

Using these equations and (3) we obtain, as we should, the equation of continuity for the complete (neutral) gas, viz.,

$$\text{div } (N\mathbf{W}) = -\partial N/\partial t \tag{8}$$

In the passage of a small wave disturbance through the steadily streaming neutral material $\mathbf{V} - \mathbf{U}$, $\hat{\mathbf{V}} - \mathbf{U}$ will be small and so, neglecting second-order small quantities we can take

$$D/Dt \doteq \partial/\partial t + (\mathbf{U} \,.\, \text{grad}) \tag{9}$$

Equations (4) and (5), when used in conjunction with (3) then yield an expression for $D\mathbf{W}/dt$ and then, eliminating this quantity by (1) we can prove

$$\Pi = p + \hat{p} + \frac{Ne}{C}\int^{(\mathbf{r})} (\mathbf{V} - \hat{\mathbf{V}}) \wedge \mathbf{B}_0 \, d\mathbf{r} \tag{10}$$

so that DALTON'S law of partial pressures applies in the waves to a first approximation.

We have now to apply this theory to the passage of such a small disturbance; and the equations governing the passage of the electromagnetic wave through the medium will be obtained by a perturbation process; i.e. by comparing (4) and (5) with the corresponding equations for the steady state. Since we are neglecting the effect of the internal gravitational pulls of the gas on itself then, in this perturbation, $\delta\mathbf{F} = 0$. Also, writing

$$\mathbf{V} = \mathbf{U} + \mathbf{v} \tag{11}$$

$$\mathbf{B} = \mathbf{H}_0 + \mathbf{H} \tag{12}$$

(where we are taking the permeability as unity and $\mathbf{H}$ as the magnetic field due to the wave) then $\mathbf{H}_0$ is the external (applied) magnetic field (effectively the earth's field) and $\mathbf{v}$, $\mathbf{H}$ are small quantities. Taking suffix zero to refer to steady state values and retaining only first-order small quantities, we have (for the negatively charged gas)

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{U} \,.\, \text{grad})\mathbf{v} = -\frac{e}{m}E - \nu\mathbf{v} - \frac{e}{mC}\mathbf{v} \wedge \mathbf{H}_0 - \frac{e}{mC}\mathbf{U} \wedge \mathbf{H} + \\ + \frac{\delta N}{mN_0{}^2}\,\text{grad}\, p_0 - \frac{1}{mN_0}\,\text{grad}\,(\partial p) \tag{13}$$

and

$$\partial(\delta N)/\partial t + \text{div}\,\{(\delta N)\mathbf{U}\} + \text{div}\,(N_0\mathbf{v}) = 0 \tag{14}$$

The current vector $\mathbf{I}$ in the electron gas due to the wave is

$$\begin{aligned}\mathbf{I} &= -(eN\mathbf{V} - eN_0\mathbf{U}) \\ &= -e\{(\delta N)\mathbf{U} + N_0\mathbf{v}\}\end{aligned} \tag{15}$$

so that

$$\operatorname{div}\mathbf{I} = e\,\partial(\delta N)/\partial t \tag{16}$$

Let

$$\beta^2 = (dp/d\rho)_0 \tag{17}$$

p being taken as a function of ρ (it would probably be reasonable to take this as the adiabatic relation, as in sound-wave theory). Then, neglecting second and higher order small quantities, we have

$$\delta p = \beta^2\,\delta p = \beta^2 m\,\delta N \tag{18}$$

and the first of the above perturbed equations becomes

$$\dot{\mathbf{v}} + (\mathbf{U}\,.\,\operatorname{grad})\mathbf{v} = -\frac{e}{m}\mathbf{E} - \nu\mathbf{v} - \frac{e}{mC}\mathbf{v}\wedge\mathbf{H}_0 - \frac{e}{mC}\mathbf{U}\wedge\mathbf{H} + \\ + \frac{\delta N}{mN_0^2}\operatorname{grad} p_0 - \frac{\beta^2}{N_0}\operatorname{grad}(\delta N) \tag{19}$$

The term grad p_0 depends upon the body forces $\mathbf{F}$ which hold the material in the steady state. Writing into equation (4) steady state values we have

$$0 = -\frac{e}{mC}\mathbf{U}\wedge\mathbf{H}_0 - \frac{1}{mN_0}\operatorname{grad} p_0 + \mathbf{F} \tag{20}$$

Hence equation (19) is, effectively*

$$\dot{\mathbf{v}} + (\mathbf{U}\,.\,\operatorname{grad})\mathbf{v} = -\frac{e}{m}E - \nu\mathbf{v} - \frac{e}{mC}\mathbf{v}\wedge\mathbf{H}_0 - \frac{e}{mC}\mathbf{U}\wedge\mathbf{H} + \\ + \frac{\delta N}{N_0}\left(\mathbf{g} - \frac{e}{mC}\mathbf{U}\wedge\mathbf{H}\right) - \frac{\beta^2}{N_0}\operatorname{grad}(\delta N) \tag{21}$$

2.2. *Specialization for sound waves only*

Writing the electrical terms and $\mathbf{U}$ equal to zero and using the equation of continuity we may verify that, in this special case

$$\partial^2(\delta N)/\partial t^2 + \nu\,\partial(\delta N)/\partial t + \mathbf{g}\,.\,\operatorname{grad}(\delta N) = \beta^2\operatorname{div}\operatorname{grad}(\delta N) \tag{22}$$

If the frictional and gravitational terms are neglected in this it becomes the usual scalar sound-wave equation. It may readily be verified that this approximation may be made if the frequency makes $\omega \gg \mathbf{g}/\beta$ and $\omega \gg \nu$. We shall make this approximation in our application of the general theory which follows.

* Writing $\mathbf{F} \fallingdotseq \mathbf{g}$.

2.3. *Determination of the current vectors*

We concentrate first on the electron gas and its current vector $\mathbf{I}$. After a certain amount of reduction we find from these equations

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\,.\,\text{grad} + \nu\right)\{\dot{\mathbf{I}} + (\text{div}\,\mathbf{I})\mathbf{U}\}$$
$$= \frac{N_0 e^2}{m}\dot{\mathbf{E}} - \frac{e}{mC}\dot{\mathbf{I}} \wedge \mathbf{H}_0 + \frac{N_0 e^2}{mC}\mathbf{U} \wedge \dot{\mathbf{H}} - (\text{div}\,\mathbf{I})\mathbf{g} + \beta^2\,\text{grad div}\,\mathbf{I} \quad (23)$$

Writing

$$\omega_0{}^2 = 4\pi N_0 e^2/m, \qquad \omega_H = e\,|H_0|\,/mC \quad (24)$$

$\mathbf{n}$ equal to a unit vector along $\mathbf{H}_0$ and taking Oz along the direction of $\mathbf{U}$ this is

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\frac{\partial}{\partial z} + \nu\right)\{\dot{\mathbf{I}} + (\text{div}\,\mathbf{I})\mathbf{U}\}$$
$$= \frac{1}{4\pi}\omega_0{}^2\frac{\partial}{\partial t}\left(\mathbf{E} + \frac{1}{C}\mathbf{U} \wedge \mathbf{H}\right) - \omega_H\dot{\mathbf{I}} \wedge \mathbf{n} - (\text{div}\,\mathbf{I})\mathbf{g} + \beta^2\,\text{grad div}\,\mathbf{I} \quad (25)$$

Writing all vectors proportional to $\exp\left(\frac{i\omega}{C}\sqrt{\kappa}z - i\omega t\right)$ (appropriate for propagation along columns) we have

$$\left(-i\omega + \frac{i\omega}{C}\sqrt{\kappa}U + \nu\right)\{-i\omega\mathbf{I} + (\text{div}\,\mathbf{I})\mathbf{U}\}$$
$$= -\frac{1}{4\pi}i\omega\omega_0{}^2\left(\mathbf{E} + \frac{1}{C}\mathbf{U} \wedge \mathbf{H}\right) + i\omega\omega_H\mathbf{I} \wedge \mathbf{n} - (\text{div}\,\mathbf{I})\mathbf{g} + \beta^2\,\text{grad div}\,\mathbf{I} \quad (26)$$

Similarly, the equation for $\hat{\mathbf{I}}$, the current vector in the positively charged gas, will be

$$\left(-i\omega + \frac{i\omega}{C}\sqrt{\kappa}\,U + \nu\right)\{-i\omega\hat{\mathbf{I}} + (\text{div}\,\hat{\mathbf{I}})\mathbf{U}\}$$
$$= -\frac{1}{4\pi}i\omega\hat{\omega}_0{}^2\left(\mathbf{E} + \frac{1}{C}\mathbf{U} \wedge \mathbf{H}\right) - i\omega\hat{\omega}_H\hat{\mathbf{I}} \wedge \mathbf{n} - (\text{div}\,\hat{\mathbf{I}})\mathbf{g} + \hat{\beta}^2\,\text{grad div}\,\hat{\mathbf{I}} \quad (27)$$

To obtain the equations for the current vectors in the stationary gas, say $\bar{\hat{\mathbf{I}}}$ for the positively charged part and $\bar{\mathbf{I}}$ for the negatively charged part, we can write $\mathbf{U} = 0$ in the above and make the necessary alterations: we thus have

$$(i\omega - \nu)i\omega\bar{\mathbf{I}} = -\frac{1}{4\pi}i\omega\bar{\omega}_0{}^2\mathbf{E} + i\omega\omega_H\bar{\mathbf{I}} \wedge \mathbf{n} - (\text{div}\,\bar{\mathbf{I}})\mathbf{g} + \bar{\beta}^2\,\text{grad div}\,\bar{\mathbf{I}} \quad (28)$$

$$(i\omega - \nu)i\omega\bar{\hat{\mathbf{I}}} = -\frac{1}{4\pi}i\omega\bar{\hat{\omega}}_0{}^2\mathbf{E} - i\omega\hat{\omega}_H\bar{\hat{\mathbf{I}}} \wedge \mathbf{n} - (\text{div}\,\bar{\hat{\mathbf{I}}})\mathbf{g} + \bar{\hat{\beta}}^2\,\text{grad div}\,\bar{\hat{\mathbf{I}}} \quad (29)$$

If we neglect the pressural and body force terms in equation (28) we obtain an equation which, when solved for $\bar{\mathbf{I}}$ and used in (34) (with $\mathscr{I}$ therein taken approximately equal to $\bar{\mathbf{I}}$), embodies the results of the classical magneto-ionic theory for waves through stationary media.

VOL. 17 1959/6

Note on the correction for internal gravitation

If this is allowed for, then the perturbation $\delta\mathbf{F}$ is no longer zero but satisfies the relation*

$$\text{div}\ (\delta\mathbf{F}) = 4\pi\gamma(m + \hat{m})(\delta N + \delta\bar{N}) \tag{30}$$

where γ is the gravitation constant. It is now possible to deduce the rather elegant result that the required correction will be made on adding to the right sides of (25–29) the terms $-4\pi\gamma(\rho_0,\ \rho_0)(\mathbf{I} + \bar{\mathbf{I}})$ where ρ_0 is the undisturbed mass density of the streaming gas and ρ_0 is the undisturbed density of the stationary gas, either factor being used as appropriate.

2.4. *Derivation of the equation connecting* $\mathbf{E}$ *with* $\mathscr{I}$

MAXWELL'S field equations are

$$\text{curl}\ \mathbf{H} = \frac{4\pi}{C}\mathscr{I} + \frac{1}{C}\dot{\mathbf{D}} \tag{31}$$

$$\text{curl}\ \mathbf{E} = -\frac{1}{C}\dot{\mathbf{B}} \tag{32}$$

together with

$$\text{div}\ \mathbf{D} = 4\pi\rho, \qquad \text{div}\ \mathbf{B} = 0 \tag{33}$$

From these equations, taking (as usual) the magnetic permeability as unity and also $\mathbf{D} = \mathbf{E}$ in the conducting gas (since the electrons, etc., are taken as vibrating *in vacuo*) and, writing, as above, all vectors proportional to $\exp(-i\omega t)$, we obtain

$$\text{curl curl}\ \mathbf{E} - \frac{\omega^2}{C^2}\mathbf{E} = \frac{4\pi i\omega}{C^2}\mathscr{I} \tag{34}$$

and connexion with the preceding equations is made through the equation for the total current vector

$$\mathscr{I} = \hat{\mathbf{I}} + \mathbf{I} + \hat{\mathbf{I}} + \hat{\bar{\mathbf{I}}} \tag{35}$$

The problem is now theoretically solved, for if the equations (26)–(29) could be solved for the four constituent current vectors we should, by use of (35) in (34) and by use of the relation

$$\mathbf{H} = \frac{iC}{\omega}\,\text{curl}\,\mathbf{E} \tag{36}$$

to eliminate $\mathbf{H}$, obtain an equation in the electric field $\mathbf{E}$ alone.

The practical difficulties are obviously immense and we shall therefore proceed tentatively. We shall begin by investigating the propagation of plane waves along the Oz axis (direction of $\mathbf{U}$) in the special case when $\mathbf{H}_0$ is also along this direction. This is the case which is of importance in whistling atmospherics. We shall then go on to examine propagation along columns, both simple stationary ones and the compound streaming types. In this connexion it is very interesting to remark that, for a stationary column, when the effect of the positively charged particles are neglected, it is possible to solve equation (28) for $\bar{\mathbf{I}}$. For in that case div $\bar{\mathbf{I}}$ can be expressed directly in terms of div $\mathbf{E}$ from (34), on taking $\mathscr{I}$ as effectively $\bar{\mathbf{I}}$.

* If stationary and streaming gases exist together.

3. The Propagation of Plane Waves*

We suppose that the frequency is considerably greater than the proton gyro-frequency and is also not near the proton plasma-frequency. In that case, we are neglecting $\bar{\mathbf{I}}$ and $\hat{\bar{\mathbf{I}}}$ in comparison with $\mathbf{I}$ and $\bar{\mathbf{I}}$. Expressing our vector equations in components, with all vectors varying as $\exp\left(\frac{i\omega}{C}\sqrt{\kappa}z - i\omega t\right)$ we find that two different types of waves which are quite independent of each other, can be propagated. These are:

3.1. *The "sound wave" type* ($E_x = 0$, $E_y = 0$, $E_z \neq 0$)

The propagation constant κ satisfies the equation

$$\omega_0^2 \Big/ \left\{\left(1 - \frac{U}{C}\sqrt{\kappa}\right)^2 - \frac{\beta^2}{C^2}\kappa\right\} + \bar{\omega}_0^2 \Big/ \left(1 - \frac{\beta^2}{C^2}\kappa\right) = \omega^2 \tag{37}$$

3.2. *The "electromagnetic" type* ($E_z = 0$)

The propagation constant satisfies the equation

$$\kappa - 1 - \frac{\bar{\omega}_0^2}{\omega(\omega_H - \omega)} = \left\{\left(1 - \frac{U}{C}\sqrt{\kappa}\right)\omega_0^2/\omega^2\right\} \Big/ \left(\frac{\omega_H - \omega}{\omega} + \frac{U}{C}\sqrt{\kappa}\right) \tag{38}$$

Having whistlers in mind we have taken the solution that reduces to

$$\kappa = 1 + \frac{\bar{\omega}_0^2 + \omega_0^2}{\omega(\omega_H - \omega)} \qquad \textit{if} \qquad U \to 0$$

From this equation it appears that, if the direction of propagation is against the stream velocity ($U < 0$) there is always one negative solution for $\sqrt{\kappa}$ (backward-travelling wave) and *either* two positive solutions (forward-travelling waves) *or* two conjugate complex solutions. If $U > 0$ (propagation *with* the stream) there is always one forward-travelling wave ($\sqrt{\kappa} > 0$) whereas there may, or may not, exist two backward-travelling waves. In this discussion we will confine ourselves to the question as to whether there are any circumstances under which forward-travelling waves could go out of existence. We will therefore confine ourselves to the case of propagation *against* the stream. The two positive roots for $\sqrt{\kappa}$ from (38) disappear if

$$4(1 - b\varepsilon_1)^3(a^2 + b)^2\alpha^4 - b^3(a^2 + b)\{8 + 4\varepsilon_1(5b + 9) - \varepsilon_1^2(b^2 - 18b - 27)\}\alpha^2 + \\ + 4b^6(1 + \varepsilon_1) = 0 \tag{39}$$

where

$$\alpha = U/C; \qquad b = (\omega_H - \omega)/\omega; \qquad a = \bar{\omega}_0/\omega$$
$$\varepsilon = \omega_0^2/\bar{\omega}_0^2 (= N_0/\bar{N}_0)$$
$$\varepsilon_1 = \left(1 + \frac{b}{a^2}\right)^{-1}\varepsilon \tag{39A}$$

* We are neglecting ν and the term in $\mathbf{g}$ in this theory.

VOL. 17 1959/6

Writing

$$\kappa_0 = 1 + \frac{a^2}{b} \tag{40}$$

(the value of κ when the moving medium is away)
so that

$$\varepsilon_1 = (1 - \kappa_0^{-1})\varepsilon \tag{41}$$

the above condition for the two positive roots for $\sqrt{\kappa}$ to disappear in the case $U < 0$ is equivalent to

$$\left\{\frac{8 + 4\varepsilon_1(5b + 9) - \varepsilon_1^2(b^2 - 18b - 27) - [(b + 9)\varepsilon_1 + 8]^{3/2}\sqrt{[\varepsilon_1(b + 1)]}}{8(1 - b\varepsilon_1)^3}\right\}^{1/2}$$

$$< \left|\frac{\alpha\sqrt{\kappa_0}}{b}\right|$$

$$< \left\{\frac{8 + 4\varepsilon_1(5b + 9) - \varepsilon_1^2(b^2 - 18b - 27) + [(b + 9)\varepsilon_1 + 8]^{3/2}\sqrt{[\varepsilon_1(b + 1)]}}{8(1 - b\varepsilon_1)^3}\right\}^{1/2} \times$$

$$\{b\varepsilon_1 < 1\} \tag{42}$$

$$\left|\frac{\alpha}{b}\sqrt{\kappa_0}\right| > \frac{2}{3\sqrt{3}} \qquad (b\varepsilon_1 = 1) \tag{43}$$

$$\left|\frac{\alpha}{b}\sqrt{\kappa_0}\right| >$$

$$> \left\{\frac{[(b + 9)\varepsilon_1 + 8]^{3/2}\sqrt{[\varepsilon_1(b + 1)]} - [8 + 4\varepsilon_1(5b + 9) - (b^2 - 18b - 27)\varepsilon_1^2]}{8(b\varepsilon_1 - 1)^3}\right\} \times$$

$$(b\varepsilon_1 > 1) \tag{44}$$

The case of most interest practically is that when $\omega < \omega_H$; $\varepsilon \ll 1$ $a^2 \gg b$: there is then some simplification in the above which shows that there is then a narrow band of values for $|\alpha|$ for which the two positive values of $\sqrt{\kappa}$ representing the two forward travelling waves, become complex, except, possibly, when ω/ω_H becomes so small that $\varepsilon\omega_H/\omega$ is no longer small. For when $\varepsilon\omega_H/\omega$ becomes large the condition becomes $|\alpha| > \varepsilon^{-1/2}(\omega_H/2\bar{\omega}_0)$ which almost certainly could not be satisfied in practice, owing to the smallness of α.

The question as to the physical meaning of these complex roots now arises, because, being complex numbers, one represents a disturbance which rapidly grows as z increases until the point is reached when this first-order theory fails. The writer has had considerable discussion of this with colleagues and there seems to be an element of uncertainty as to whether this "growing wave" can really exist. The writer thinks not, for the reason which follow.

If a physical disturbance of the ionized gas corresponding to this particular complex value of $\sqrt{\kappa}$ does exist we can say:

(i) it is no longer a simple travelling wave as the ion disturbances no longer obey a law of the form:

$$\exp\left\{\frac{2\pi i}{\lambda}(z - Vt)\right\}$$

(ii) the particle vibrations are very large compared with what they would be for a real value of $\sqrt{\kappa}$ (for if not first-order wave theory approximations would apply).

As we are neglecting ν and as the rate of working of the external magnetic force on a moving particle is zero (ii) would imply a relatively large storing of energy by the medium. Since (for $U < 0$) there is a negative solution for $\sqrt{\kappa}$, and since the solution corresponding to (38) which we have so far left out, viz.

$$\kappa - 1 + \frac{\bar{\omega}_0{}^2}{\omega(\omega_H + \omega)} = -\left\{\left(1 - \frac{U}{C}\sqrt{\kappa}\right)\omega_0{}^2/\omega^2\right\}\Big/\left(\frac{\omega_H + \omega}{\omega} - \frac{U}{C}\sqrt{\kappa}\right) \quad (45)$$

gives, for all negative U, just one real value of $\sqrt{\kappa}$, which is also negative it seems clear to the writer that the case under consideration corresponds to reflection of any energy which might originally have been travelling in the Oz direction, via either or both of the backward travelling waves (which are given by the above negative permitted values of $\sqrt{\kappa}$). This interpretation is quite consistent with the usual way in which "reflection at a thick uniformly ionized layer at frequencies above the critical" is arrived at. For such a layer, ignoring the effect of the earth's field

$$\begin{aligned} z > 0, \qquad & \kappa = 1 - \frac{\omega_0{}^2}{\omega(\omega + i\nu)} \; (\nu \ll \omega) \\ z < 0, \qquad & \kappa \equiv 1 \end{aligned} \quad (46)$$

The case $\omega > \omega_0$ gives complex values of $\sqrt{\kappa}$ (which are similar in type to those under discussion above, being not strictly pure imaginary): we have also a negative real root $\sqrt{\kappa} = -1$ to provide a channel for the carrying of reflected energy. The universally accepted interpretation here is to say that the reflection is total with the forward part in $z > 0$ exponentially damped, the "growing wave" corresponding to the other complex root being rejected. It is therefore only consistent to adopt a similar interpretation above.

Doing this, it now follows from the above theory that, under suitable circumstances, wave propagation against a streaming gas can suffer sharp breaks as the frequency is varied (U/C fixed). See (42), for example, which, when $\varepsilon\omega_H/\omega \ll 1$, is approximately equivalent to (using $a^2 \gg b$)

$$1 - \frac{\sqrt{2}}{8}(\varepsilon\omega_H/\omega)^{1/2} < \left|\frac{U}{C}\right| \frac{\bar{\omega}\sqrt{\omega}}{(\omega_H - \omega)^{3/2}} < 1 + \frac{\sqrt{2}}{8}(\varepsilon\omega_H/\omega)^{1/2}$$

Since records of whistling atmospherics exhibiting abrupt truncations are sometimes observed the above may provide a possible theoretical explanation of the phenomenon.

Acknowledgements—It is a pleasure to acknowledge here my indebtedness to L. R. O. Storey of the National Bureau of Standards, U.S.A., who suggested the investigation, to C. Hines and R. Barrington of the Defence Research Board of

VOL. 17 1959/6

Canada for helpful discussions and suggestions and to M. KLINE and B. FRIEDMAN of the Division of Electromagnetic Research, New York University, for some useful discussions.

REFERENCES

HELLIWELL R. A.	1953–1956	*Low Frequency Propagation Studies* Part 1. Final Report AFCRC-TR-189. ASTIA document AD 110184.
SMITH R. L.	1957	The Guiding of Whistlers. *Proceedings of the Boulder Symposium on the Propagation of VLF Radio Waves.*
STOREY L. R. O.	1958	*A Method to Interpret the Dispersion Curves of Whistlers.* Report No. 23-4-1. Radio Physics Laboratory, Defence Research Board Telecommunications Establishment, Ottawa, Canada.

Note added in proof

The writer has recently succeeded in solving equations (26) and (27) for $\mathbf{I}$ and $\hat{\mathbf{I}}$ respectively, neglecting the pressural and gravitational terms. This has made it possible to discuss the case of electromagnetic wave propagation in the compound streaming media. This work (with an appraisal of the conditions under which the above-mentioned approximation is valid) will appear shortly in the Proceedings of the 3rd U.R.S.I. Symposium on Electromagnetic Theory held at the University of Toronto, Canada, June 1959. These Proceedings are being published soon by the American Institute of Radio Engineers as a special issue under their Professional Group on Antennae and Propagation.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 170 to 178. Pergamon Press Ltd. Printed in Northern Ireland

The propagation of electromagnetic waves in ionized gases (with special reference to "whistlers")—II*

F. H. NORTHOVER
Department of Mathematics, Carleton University, Ottawa, Canada

(*Received* 1 *June* 1959)

1. THE EQUATION FOR E

THE general equation for the electric field $\mathbf{E}$ of an electromagnetic wave propagated through a stationary ionic distribution will be obtained from equation (34) of Part I, $\bar{\mathbf{I}}$ and $\hat{\bar{\mathbf{I}}}$ being determined by equations (28) and (29) of that paper.

In what follows we make the assumption that the wave is almost entirely carried by the negatively charged particles: this is equivalent to assuming that the wave-frequency is large compared with the proton gyro-frequency (cf. Appendix). Consequently, we take $\mathscr{I}$ as effectively equal to $\bar{\mathbf{I}}$ which by (34) of Part I implies

$$\operatorname{div} \bar{\mathbf{I}} = \frac{i\omega}{4\pi} \operatorname{div} \mathbf{E} \tag{1}$$

so that (28) of Part I becomes

$$(i\omega - \nu)\hat{\bar{\mathbf{I}}} = -\frac{1}{4\pi}\bar{\omega}_0{}^2\,\mathbf{E} + \omega_H \bar{\mathbf{I}} \wedge \mathbf{n} - \frac{1}{4\pi}\{(\operatorname{div}\mathbf{E})\mathbf{g} - \bar{\beta}^2 \operatorname{grad}\operatorname{div}\mathbf{E}\} \tag{2}$$

This equation can be solved for $\bar{\mathbf{I}}$ and then, taking $\mathscr{I} \doteq \bar{\mathbf{I}}$ again in (34), Part I, this yields an equation in $\mathbf{E}$ alone. At the frequencies we are considering the term in $\mathbf{g}$ above is unimportant: we shall therefore neglect it (cf. equation 22, Part I, *et seq.*). Equation (2) therefore becomes

$$(i\omega - \nu)\,\bar{\mathbf{I}} = -\frac{1}{4\pi}\,\bar{\omega}_0{}^2\,\bar{\mathbf{A}} + \omega_H \bar{\mathbf{I}} \wedge \mathbf{n} \tag{3}$$

where

$$\bar{\mathbf{A}} = \mathbf{E} - \frac{\bar{\beta}^2}{\omega_0{}^2} \operatorname{grad}\operatorname{div}\mathbf{E} \tag{4}$$

The solution of this vector equation is

$$\bar{\mathbf{I}} = \frac{i\omega}{4\pi}\,\bar{\alpha}\,\{\bar{\mathbf{A}} - \xi^2(\mathbf{n}\cdot\bar{\mathbf{A}})\mathbf{n} - i\xi\,\bar{\mathbf{A}} \wedge \mathbf{n}\} \tag{5}$$

where

$$\bar{\alpha} = \bar{\sigma}^2\xi\gamma/(1 - \xi^2);\ \bar{\sigma} = \bar{\omega}_0/\omega;\ \xi = \omega_H/(\omega + i\nu);\ \gamma = \omega/\omega_H \tag{5A}$$

* Contribution from the Radio Physics Laboratory of the Defence Research Telecommunications Establishment, Defence Research Board, Ottawa, Canada, under PCC D48-95-11-01, and from the Department of Mathematics, Carleton University, Ottawa, Canada.

VOL. 17 1959/6

The equation satisfied by $\mathbf{E}$ is therefore (from equation 34, Part I)

$$\text{curl curl } \mathbf{E} + k^2\{\bar{\varkappa}[\bar{\mathbf{A}} - \xi^2(\mathbf{n}\cdot\bar{\mathbf{A}})\mathbf{n} - i\xi\bar{\mathbf{A}}\wedge\mathbf{n}] - \mathbf{E}\} = 0 \tag{6}$$

where
$$k = \omega/C \tag{6A}$$

2. Standard Type Propagation down Non-homogeneous Columns

2.1. The radial distribution of the ionic density will be taken in the general form

$$\bar{N} = \bar{N}_I + (\bar{N}_C - \bar{N}_I)f(r/a) \tag{7}$$

where r is the distance from the axis of any point of the distribution, $f(0) = 1$, $f'(0) = 0$, $f' < 0$, otherwise: $f(\infty) = 0$. $\bar{N}_C$ is the value of $\bar{N}$ at the axis and $\bar{N} \to \bar{N}_I$ as $r \to \infty$.

The length a will be a measure of the "extent" or "thickness" of the irregularity: as $a \to \infty$, $d\bar{N}/dr \to 0$ for every r. Hence, according to the definition of standard type propagation (cf. Part I) we require a solution for $\mathbf{E}$ which approaches the simple plane travelling wave (the components being cartesian)

$$\mathbf{E}_\infty = (1, i, 0)\exp\left(\frac{i\omega}{C}\sqrt{\kappa}z - i\omega t\right) \tag{8}$$

as $a \to \infty$, where, since $\lim_{a\to\infty} \bar{N} = \bar{N}_C$, κ takes the value appropriate to plane wave propagation in a homogeneous medium $\bar{N} = \bar{N}_C$; viz.

$$\kappa = 1 - \bar{\alpha}_C - \bar{\alpha}_C\xi \tag{9}$$

the suffix c denoting the value of $\bar{\alpha}$ when $\bar{N} = \bar{N}_c$.

The standard type of solution to seek, then, for "standard type" propagation, is

$$\mathbf{E} = \mathbf{F}\exp\left(\frac{i\omega}{C}\sqrt{\kappa}z + i\varphi - i\omega t\right) \tag{10}$$

where the cylindrical components of the vector $\mathbf{F}$ are functions of r *only* which tend to 1, i, 0, respectively, when $a \to \infty$; and κ takes the value (9).

In classical magneto-ionic theory the pressure term (i.e. that involving $\bar{\beta}$) is usually neglected. This does not cause any error when only plane waves parallel to the applied magnetic field in homogeneous media are considered, as we saw in Part I, Section 3.

When propagation down columns is considered, however, the term in $\bar{\beta}$ enters. Owing to the complexity which is introduced by retention of this term, however, we shall proceed tentatively: first, we shall discuss the propagation when the pressure term is neglected, second, we shall try to elucidate the circumstances under which this approximation can be made, and, last, we shall try to extend the analysis to cover the $\bar{\beta}$ term.

Neglecting, for the time being, therefore, the term in $\bar{\beta}$, i.e. taking $\mathbf{\Lambda}$ as

effectively equal to **E**. we find, on working out the components of curl **E** in terms of **F** and substituting in (6) the following equations for F_1, F_2, F_3:

$$\left\{\begin{aligned} & ik\sqrt{\kappa}F_3' + ir^{-1}F_1' + r^{-1}(F_1 + iF_2) + \\ & \qquad + k^2\{(\kappa - 1 + \bar{\varkappa})F_1 - i\bar{\varkappa}\xi F_3\} = 0 \\ & -F_2'' - r^{-1}(k\sqrt{\kappa}F_3 + F_2' - iF_1') + r^{-2}(F_2 - iF_1) + \\ & \qquad + k^2\{(\kappa - 1 + \bar{\varkappa})F_2 + i\bar{\varkappa}\xi F_1\} = 0 \\ & -F_3'' + r^{-1}(ik\sqrt{\kappa}F_1 - k\sqrt{\kappa}F_2 - F_3') + r^{-2}F_3 + ik\sqrt{\kappa}F_1' + \\ & \qquad + k^2(\bar{\sigma}^2\xi\gamma - 1)F_3 = 0 \end{aligned}\right\} \tag{11}$$

the dash denoting d/dr.

The distribution is expressed as a function of r/a, not r, and we have to satisfy the conditions $\mathbf{F} \to (1, i, 0)$ as $a \to \infty$, for every fixed r. Writing therefore,

$$\tau = r/a \tag{12}$$

and also

$$\bar{\varkappa} = \bar{\varkappa}_c - \Delta\bar{\varkappa}, \qquad \bar{\sigma} = \bar{\sigma}_c^{\,2} + \delta \tag{12A}$$

the equations become

$$\left\{\begin{aligned} & ika\sqrt{\kappa}F_3' + i\tau^{-1}F_2' + \tau^{-2}(F_1 + iF_2) + \\ & \qquad + k^2a^2\{\bar{\varkappa}_c\xi(-F_1 - iF_2) - \Delta\bar{\varkappa}(F_1 - i\xi F_2)\} = 0 \\ & -F_2'' - \tau^{-1}(ka\sqrt{\kappa}F_3 + F_2' - iF_1) + \tau^{-2}(F_2 - iF_1) + \\ & \qquad + k^2a^2\{\bar{\varkappa}\xi(-F_2 + iF_1) - \Delta\bar{\varkappa}(F_2 + i\xi F_1)\} = 0 \\ & -F_3'' + \tau^{-1}(ika\sqrt{\kappa}F_1 - ka\sqrt{\kappa}F_2 - F_3') + \tau^{-2}F_3 + ika\sqrt{\kappa}F_1' + \\ & \qquad + k^2a^2(\bar{\sigma}_c^{\,2}\xi\gamma - 1 + \xi\gamma\delta)F_3 = 0 \end{aligned}\right\} \tag{13}$$

where the dash now denotes $d/d\tau$.

2.2. *The field at long distances from the axis*

When τ is large enough we can find an asymptotic approximation of the type

$$\mathbf{F} = \exp(iKka\tau) \sum_{n=0} \mathbf{a}_n \tau^{u-n} \tag{14}$$

it being assumed that $f(\tau)$ is expansible in powers of $1/\tau$. Writing $\mathbf{a}_n = (A_n, B_n, C_n)$ we have from the highest order terms,

$$\left\{\begin{aligned} -K\sqrt{\kappa}C_0 - (\bar{\varkappa}\xi + \bar{\Delta})A_0 - (\bar{\varkappa} - \bar{\Delta})i\xi B_0 = 0 \\ \{K^2 - (\bar{\varkappa}_c\xi + \bar{\Delta})\}B_0 + (\bar{\varkappa}_c - \bar{\Delta})i\xi A_0 = 0 \\ (K^2 + \bar{\sigma}_c^{\,2}\xi\gamma - 1)C_0 - K\sqrt{\kappa}A_0 = 0 \end{aligned}\right\} \tag{15}$$

where $\bar{\Delta}$ denotes the value of $\Delta\bar{\varkappa}$ far from the axis.

Therefore K satisfies the quartic equation

$$\begin{aligned} & K^4(\kappa + \bar{\varkappa}_c\xi + \bar{\Delta}) - K^2[(\bar{\Delta} + \bar{\varkappa}_c\xi)(\kappa + 1 - \bar{\sigma}_I^{\,2}\xi\gamma) + \bar{\Delta}(1 + \xi) \times \\ & \qquad \{2\bar{\varkappa}_c\xi + \bar{\Delta}(1 - \xi)\}] - (\bar{\sigma}_I^{\,2}\xi\gamma - 1)(1 + \xi)\{2\bar{\varkappa}_c\xi + \bar{\Delta}(1 - \xi)\}\bar{\Delta} = 0 \end{aligned} \tag{16}$$

By considering in detail the equations for the higher coefficients (A_i, B_i, C_i),

VOL. 17 1959/6

$i \geqslant 1$ it appears that $u = -\frac{1}{2}$ (as might have been expected from energy considerations) and, clearly, only roots of this quartic in K for which Im $(K) \geqslant 0$ are physically admissible.

The asymptotic approximation is therefore of the form

$$\mathbf{F} \sim \exp(iK_1 ka\tau) \sum_{n=0} \mathbf{a}_n^{(1)} \tau^{-1/2-n} + \exp(iK_2 ka\tau) \sum_{n=0} \mathbf{a}_n^{(2)} \tau^{-1/2-n} \tag{17}$$

where K_1 and K_2 are the two allowable roots of the quartic.

Substituting the values of the parameters in terms of the various frequencies and neglecting the collision frequency, we have

$$\begin{aligned} &\{1 - \bar{\Lambda} + \bar{G}(1 - \gamma^2)\}K^4 + \frac{1}{G\gamma^3(1-\gamma)} \{2\gamma - 1 + \bar{\Lambda}(2\gamma^2 - 4\gamma + 1) + \\ &\quad + 2\bar{\Lambda}^2\gamma(1-\gamma) + 2\bar{G}\gamma^2(1-\gamma)(1 + \bar{\Lambda}\gamma)\}K^2 - \\ &\quad - \frac{1}{\bar{G}^2\gamma^4(1-\gamma)} (1 - \bar{\Lambda} - \bar{G}\gamma^2)\{2 - \bar{\Lambda}(1-\gamma)\}\bar{\Lambda} = 0 \end{aligned} \tag{18}$$

where

$$\begin{aligned} \bar{\Lambda} &= (\bar{N}_c - \bar{N}_I)/\bar{N}_c \\ \gamma &= \omega/\omega_H \quad \text{(as already defined)} \\ \bar{G} &= \omega_H{}^2/\bar{\omega}_c{}^2 \end{aligned} \tag{19}$$

For strong guiding of the waves by the column we require that:

(1) The field should decrease radially at a much greater rate than does the radial scattering factor $r^{-(1/2)}$.

(2) With condition (1) satisfied the central regions of the column should carry a reasonably high proportion of the total energy carried by the cylindrical distribution.

We are now in a position to examine the circumstances under which the fundamental requirement (1) holds. To elucidate condition (2) we need to know the law of distribution of the column which (for example) affects the constant coefficients A_i, B_i, C_i, in the above asymptotic expression. It is therefore much more difficult to discuss than (1). Owing to the length and complexity of the analysis necessary to deal properly with this point only a summary of the position will be given in this paper: full details will appear elsewhere.*

For condition (1) to be satisfied we require *either* that $K_1{}^2$ and $K_2{}^2$ be both negative, *or* that they be both complex.

(*Note*: In connexion with condition (1), $\bar{\Lambda}$ cannot, of course, be allowed to become too small, because, since one of the roots K_1, K_2 tends to zero with $\bar{\Lambda}$, r would have to become very large indeed before the radial exponential decline stipulated by condition (1) made itself felt. This limitation on $\bar{\Lambda}$ would appear in the full discussion.)

We note that, essentially, $\bar{\Lambda} < 1$ and, in the cases in which we are interested, $\gamma < 1$. Bearing this in mind we find, from consideration of equation (18) the following requirements for the satisfaction of condition (1):

* The Symposium Proceedings: see references.

(a) When $\bar{N}_c > \bar{N}_I$ ($\bar{\Lambda} > 0$) the necessary and sufficient condition that the field should decline exponentially in a radial direction outwards from the column at large distances from the axis is

$$\left.\begin{array}{ll} \text{(i)} & \bar{N}_I/\bar{N}_c < f_H{}^2/f_c{}^2 \\ \text{(ii)} & (\bar{N}_I/\bar{N}_c)^{1/2}(f_c/f_H) < \gamma < 1 \end{array}\right\} \tag{20}$$

where f_H, f_c are, respectively, the gyro-frequency and the axial plasma-frequency. The cut-off frequency is here $(\bar{N}_I/\bar{N}_c)^{1/2} f_c$ so that for this guiding effect to hold over most of the frequency range $0 < f < f_H$, we must have

$$(\bar{N}_I/\bar{N}_c)^{1/2}(f_c/f_H) \ll 1 \tag{21}$$

(b) When $\bar{N}_c < \bar{N}_I$ ($\bar{\Lambda} < 0$) the necessary and sufficient condition that the field should decline exponentially in a radial direction outwards from the column at large distances from the axis is

$$\gamma_1 < \gamma < \min\{1;\ (\bar{N}_I/\bar{N}_c)^{1/2}(f_c/f_H)\} \tag{22}$$

where γ_1 is the root between 0 and 1 of the equation

$$1 + 2\left(\frac{\bar{G}\gamma^2 - 1}{1 - \bar{\Lambda}}\right)(1 - \gamma)(1 + \bar{\Lambda}\gamma) = 0 \tag{23}$$

which gives

$$\gamma_1 \doteq 1/(-2\bar{\Lambda}) \tag{24}$$

for $-\bar{\Lambda} \gg 1$.

Hence for this guiding effect to hold over most of the frequency range (condition 1 holding) $0 < f < f_H$ we must have

$$\left\{\begin{array}{c} \text{(i)} \quad \bar{N}_I/\bar{N}_c > f_H{}^2/f_c{}^2 \\ \text{or these two ratios approximately equal and,} \\ \text{(ii)} \quad \bar{N}_I/\bar{N}_c \gg 1 \end{array}\right\} \tag{25}$$

2.3. *The field near the axis*

2.3.1. As has been pointed out already, to have a rapidly decreasing field in a radial direction at sufficient distances from the axis is a necessary but not sufficient condition for the column to act as a good localized guide of electromagnetic energy. For this we require also that the central (axial) regions of the column carry a reasonable proportion of the total energy carried by the distribution.

We need, therefore, an estimate of the field near the central part of the column. As the preceding asymptotic type approximation is inadequate for this, we shall develop a series type approximation. In order to do this, we must suppose that the distribution function $f(\tau)$ is expansible in a series of powers of τ^2, say

$$f(\tau) = 1 - f_1\tau^2 - f_2\tau^4 - \ldots \tag{26}$$

As $a \to \infty$ we must have $F_1 \to 1$, $F_2 \to i$, $F_3 \to 0$. Also, it can be shown, from consideration of the equations (13) for these quantities, that series expansions

VOL. 17 1959/6

for F_1 and F_2 can contain only even powers of τ and a series for F_3 only odd powers. We therefore assume that

$$\left.\begin{aligned} F_1 &= 1 + \lambda_1\tau^2 + \lambda_2\tau^4 + \dots \\ F_2 &= i + \mu_1\tau^2 + \mu_2\tau^4 + \dots \\ F_3 &= \nu_0\tau + \nu_1\tau^3 + \dots \end{aligned}\right\} \tag{27}$$

For any r, $\tau \to 0$ as $a \to \infty$. The coefficients in the series tend to infinity with "a" but it can be shown that they do not increase sufficiently rapidly with "a" to prevent every variable term in these series tending to zero as "a" $\to \infty$. Thus the series satisfies the fundamental requirement above on F_1, F_2, F_3.

Denoting by the abbreviation "$\bar{\Delta}$" the value of $\Delta\bar{\varkappa}$ far from the axis, then

$$\bar{\Delta} = \bar{\varkappa}_c - \bar{\varkappa}_I \tag{28}$$

and

$$\begin{aligned} \bar{\varkappa} &= \bar{\varkappa}_I + (\bar{\varkappa}_c - \bar{\varkappa}_I)f(\tau) \\ &= \bar{\varkappa}_c - \bar{\Delta}\{1 - f(\tau)\} \end{aligned}$$

so that in the equations (13) we use

$$\Delta\bar{\varkappa} = \bar{\Delta}\{1 - f(\tau)\} \tag{29}$$

where $f(\tau)$ is expanded as in (26).

Recurrence equations to determine the coefficients of (27) can be obtained (cf. the Symposium Proceedings): but we shall merely quote here the values of the first few of them and relevant general properties: we find $\nu_0 = 0$ and

$$\left\{\begin{aligned} \lambda_1 &= -\frac{3}{4}\frac{f_1(1+\xi)\bar{\Delta}}{1-\bar{\varkappa}_c} \\ \mu_1 &= -\frac{i}{4}\frac{f_1(1+\xi)\bar{\Delta}}{1-\bar{\varkappa}_c} \\ \nu_1 &= -\frac{ika\sqrt{\kappa}}{4}\frac{f_1(1+\xi)\bar{\Delta}}{1-\alpha_c} \end{aligned}\right\} \tag{30}$$

The values of λ_2, μ_2 and ν_2 have been calculated and are given in the above-mentioned reference where it is also shown that λ_n, μ_n are polynomials of degree $2n - 2$ in a, while ν_n is a polynomial of degree $2n - 1$ in a (containing at least a factor a): F_1, F_2 and F_3 therefore tend to the required limits 1, i, 0 as $a \to \infty$.

2.3.2. *The energy flow in the axial regions.* Using the series expressions, and taking into account the asymptotic behaviour of the field for large τ, it is possible to gain an estimate of the amount of energy carried by the column and also of the degree in which it is confined to the axial regions.* The analysis necessary to discuss this is, however, too lengthy to be reproduced here: it will be given separately in the above-mentioned reference. We therefore summarize the results below:

Case $\bar{\Lambda} \ll -1$; *permitted* f-*range* $\frac{1}{2}\mathrm{f_H}/(-\bar{\Lambda}) < \mathrm{f} < \mathrm{f_H}$. The energy flux down the column approximates to

$$\tfrac{1}{4}C\lambda_c^2\chi \tag{31}$$

* $\tau < \tau_m$ where ika Im $(K)\tau_m = \pi$; c.f. (17)

where C = the velocity of light

$$\left\{\begin{array}{c} 2\pi/\lambda_c = \bar{\omega}_c/C, \text{ the suffix denoting axial values} \\ \chi = \dfrac{1}{4}\dfrac{\bar{f}_c}{f_H}\dfrac{1}{(-\overline{\Lambda})\{\gamma(1-\gamma)\}^{1/2}}\ ;\ \gamma = f/f_H \end{array}\right\} \tag{31A}$$

provided that $X^2 \gg \overline{\Lambda}$ or is near $\overline{\Lambda}$, where

$$X = a/\lambda_c \tag{31B}$$

If $X^2 \ll |\overline{\Lambda}|$ it appears that χ is still of this order of magnitude, $|\overline{\Lambda}|$ being large.

For good guiding of energy, however, X cannot be allowed to become too small otherwise the field spreads appreciably away from the axial regions, as appears from consideration of the asymptotic analysis. This extra requirement is found to be

$$X > |\overline{\Lambda}|^{-1/2} \tag{32}$$

Case $0 < 1 - \overline{\Lambda} \ll 1$: *permitted* f-*range* $[(1-\overline{\Lambda})/\mathrm{G}]^{1/2}\mathrm{f_H} < \mathrm{f} < \mathrm{f_H}$. If $\gamma \gg \chi^{-1/2}$

$$\chi \doteq \gamma^{-3/2}(1-\gamma)^{1/2}(f_c/f_H) \tag{33}$$

$$= 0\{X^{-2}\gamma^{-11/2}(1-\gamma)^{5/2}(f_c/f_H)\} \tag{34}$$

if γ is near or $\ll X^{-1/2}$.

In this case the requirement that there be little spreading of energy beyond the axial regions ($\tau < \tau_m$) amounts to $\gamma > \frac{1}{4}X^{-2}$.

3. Conclusion

Energy can be carried either by columns of the central ionic surplus type (which we shall call S-columns) or by the central ionic deficiency type ("D-columns") over a large part of the low frequency range $0 < f < f_H$, provided that both types are well developed ($\bar{N}_e \ll \bar{N}_I$ or $\bar{N}_e \gg \bar{N}_I$), if only the columns are not too thin.

The main differences in the two cases is in the order of magnitude of the energy parameter χ: apart from simple factors in γ, there is an extra small factor $|\overline{\Lambda}|^{-1}$ in the expression for D-columns (cf. 31A and 33). It is therefore clear that, when the thickness parameter X is large enough for both types of column to be operative, far more energy is carried down an S-column than down a D-column. There are, however, some qualifying conditions over and above the necessity condition that the columns should be well-developed, connected, as mentioned above, with the columns' thicknesses. These are now explained.

If the S-columns are operative they are much better energy carriers than the D-columns—but they have to be rather thick [$X > \frac{1}{2}\gamma_L^{-1/2}$, where γ_L, the cut-off lower value of γ, is given by $\gamma_L f_H = (\bar{N}_I/\bar{N}_e)^{1/2}f_c$]. On the other hand, quite thin D-columns are permissible [$X > \sqrt{\gamma_L}$, where γ_L , the lower cut-off value of γ for these columns, is approximately equal to $1/(-2\overline{\Lambda})$, for $\overline{\Lambda}$ large]. However, since for this type of column, $-\overline{\Lambda}$ cannot much exceed f_c/f_H (otherwise the carried energy becomes small) it follows that γ_L for this type of column cannot be much less than $\frac{1}{2}(f_H/f_c)$. This may not allow a low enough cut-off frequency ($\gamma_L f_H$) to

VOL. 17 1959/6

enable observed propagation to be explained in terms of this type of column. In view of this, therefore, and in view of the above-noted fact that the flux of energy along S-columns is so much greater than that along D-columns, it seems likely that the propagation of whistling atmospherics can be accounted for by supposing the existence of columns of the former type.

Note on the effects of gas pressure

Extensive consideration of equation (4), in which this term appears, by means of a perturbation technique based upon the above known solution of the case $\bar{\beta} \equiv 0$ shows that neglection of this effect is justified provided that

$$\begin{aligned} &(\bar{\beta}/C)^2/(1-\gamma) \ll 1 \qquad (\overline{\Lambda} > 0) \\ &(\bar{\beta}/C)^2 \ll \min\left\{\frac{\bar{f}_0}{f} \mid \overline{\Lambda} \mid^{-1/2};\ \mid \overline{\Lambda} \mid^{-3/2}\right\} \quad (\overline{\Lambda} < 0) \end{aligned} \tag{35}$$

Acknowledgements—It is a pleasure to acknowledge here my indebtedness to L. R. O. Storey of the National Bureau of Standards, Boulder, U.S.A., who suggested the investigation; to Messrs. C. Hines and R. Barrington of the Defence Research Telecommunications Establishment, Defence Research Board, Ottawa, Canada, who suggested extension of the investigation to moving columns; to M. Kline of the Division of Electromagnetic Research, Institute of Mathematical Sciences, New York University, New York, for some useful discussions; and to R. Dobell, of the University of British Columbia, Vancouver, Canada, for checking much of the algebraical work.

OL.
17
9/60

Appendix

The Conduction Current Equation in the Neutral Ionized Gas

Neglecting the effect of the pressure terms (which seems reasonable, in view of the investigation of the text) we find from (5) of the text

$$\bar{\mathbf{I}} = \frac{i\omega}{4\pi}\,\bar{\alpha}\,\{\mathbf{E} - \xi^2(\mathbf{n}\cdot\mathbf{E})\mathbf{n} - i\xi\mathbf{E}\wedge\mathbf{n}\}$$

where

$$\bar{\alpha} = \bar{\sigma}^2\xi\gamma/(1-\xi^2);\ \bar{\sigma} = \bar{\omega}_0/\omega;\ \xi = \omega_H/(\omega + i\nu);\ \gamma = \omega/\omega_H$$

Similarly (using the notation of Part I) we have, for the proton gas

$$\hat{\bar{\mathbf{I}}} = \frac{i\omega}{4\pi}\,\hat{\bar{\alpha}}\,\{\,\mathbf{E} - \xi^2\,(\mathbf{n}\cdot\mathbf{E})\,\mathbf{n} + i\xi\mathbf{E}\wedge\mathbf{n}\,\}$$

where

$$\hat{\bar{\alpha}} = \hat{\bar{\sigma}}^2\,\hat{\xi}\,\hat{\gamma}/(\chi - \hat{\xi}^2);\ \hat{\bar{\sigma}} = \hat{\bar{\omega}}_0/\omega;\ \xi = \hat{\omega}_H/(\omega + i\hat{\nu});\ \hat{\gamma} = \hat{\omega}/\omega_H$$

Adding these two current vectors to obtain the total current we find that the equation for $\mathbf{E}$ is now

$$\operatorname{curl}\operatorname{curl}\mathbf{E} - \frac{\omega^2}{C^2}\mathbf{E} = -\frac{\omega^2}{C^2}\{(\bar{\alpha} + \hat{\bar{\alpha}})\,\mathbf{E} - (\bar{\alpha}\,\xi^2 + \hat{\bar{\alpha}}\,\hat{\xi}^2)(\mathbf{n}\cdot\mathbf{E})\mathbf{n} - i(\bar{\alpha}\xi - \hat{\bar{\alpha}}\hat{\xi})\,\mathbf{E}\wedge\mathbf{n}\}$$

Considering a plane wave with all fields proportional to exp $(ik\sqrt{\kappa}\xi)$ we find from this that

$$\kappa - 1 + \bar{\alpha}_c + \hat{\bar{\alpha}}_c = -(\bar{\alpha}_c\xi - \hat{\bar{\alpha}}\bar{\xi})$$

Hence, neglecting ν and $\hat{\nu}$, we find

$$\kappa = 1 - \frac{\bar{\omega}_c{}^2}{\omega(\omega_H - \omega)}\left\{1 - \frac{m}{\hat{m}}\frac{1-\gamma}{(m/\hat{m}) + \gamma}\right\}$$

which approximates to the value of K used in the text when $\omega \gg \hat{\omega}_H$.

References

The results which have been quoted herein have been established in the further treatment of these problems, already referred to in Part I, which will shortly appear in the Proceedings of the 3rd U.R.S.I. Symposium on Electromagnetic Theory held at the University of Toronto, Canada, June 1959. These Proceedings are being published soon by the American Institute of Radio Engineers as a special issue under their Professional Group on Antennae and Propagation.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 17, pp. 179. Pergamon Press Ltd. Printed in Northern Ireland

BOOK REVIEW

J. W. DUNGEY: **Cosmic Electrodynamics.** Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, 1958. 183 pp. 32s. 6d.

MODERN cosmological studies are frequently leading to the conclusion that magnetic fields are associated with celestial objects. The suggestion has even been made that there is some connexion between large-scale interstellar magnetic fields and the spiral arms of galaxies. It is true that the suggestion, made some years ago, that a rotating massive body inherently possesses a magnetic field has now been abandoned; and it is now considered that Maxwell's equations are in no need of supplementation, or revision.

Professor H. ALFVEN's book *Cosmical Electrodynamics*, published in 1950, courageously surveyed the cosmical problems which invited solution by way of electromagnetic theories — and pointed out that, under certain conditions, electromagnetic forces could greatly exceed gravitational forces. Dr. DUNGEY's new volume, with the same title, is a more up-to-date theoretical survey of much the same territory in so far as it deals with magnetodynamics, solar phenomena, aurorae and the ionosphere. Yet it is far more than that. It is a work of mature scholarship, notably lucid and economic in style. Its nine chapters are titled as follows: Preliminaries; The Velocity Distribution Method; The Motion of Magnetic Fields; Magnetostatics; Magnetodynamics; Acceleration of Charged Particles to High Energy; Solar Phenomena; Magnetic Storms and Aurorae; and Ionospheric Electrodynamics. Emphasis is throughout placed on fundamental aspects of the theory which, in addition to the subjects mentioned, finds application in problems arising in the development of controlled thermonuclear reactors, such as Zeta, Sceptre and the Perhapsatron.

E. V. APPLETON

OL.
17
9/60

Papers to be published in future issues

K. C. Yeh and O. G. Villard, Jr.: Fading and attenuation of high-frequency radio waves propagated over long paths crossing the auroral, temperate and equatorial zones

C. M. Minnis and G. H. Bazzard: The geographical distribution of ionization in the *F2*-layer

H. J. A. Chivers: The simultaneous observation of radio star scintillations on different radio frequencies

W. C. Bain: The azimuth distribution of oblique reflections from meteor trails and its relation to meteor radiant distributions

H. N. Rundle, D. M. Hunten and J. W. Chamberlain: Resonance scattering by atmospheric sodium—VII. Measurement of the vertical distribution in twilight

M. L. White: Thermal and gravitational atmospheric oscillations—ionospheric dynamo effects included

D. M. Schlapp: An attempt to measure the collision frequency of electrons in the *F*-region of the ionosphere

W. C. Hoffman: Conditions for the persistence of purely longitudinal or purely transverse propagation

B. Chatterjee: A note on fourth reflection condition in the ionosphere

L. R. Megill: Photometric observations of the twilight glow [OI] 5577 and [OI] 6300

A. Onwumechilli: Fluctuations in the geomagnetic horizontal field near the magnetic equator

D. M. Hunten: Resonance scattering by atmospheric sodium. An improved method of deducing the vertical distribution

D. H. McPherson and A. Vallance Jones: A study of the latitude dependence of OH rotational temperatures for Canadian stations

Y. S. N. Murty: Group refractive index of the ionosphere at low frequency

D. W. G. Chappell: A relationship between the life of M-regions and the rate of change of solar activity

J. Veldkamp: A giant geomagnetic pulsation

S. Shanack: A theoretical current density Ansatz for the quiet day solar semi-diurnal tidal mode of oscillation of the ionosphere

R. L. Dowden: Ionospheric thermal radiation at radio frequencies in the auroral zone

G. R. A. Ellis: Gyro splitting of ionospheric echoes

M. S. V. Gopala Rao, B. Ramachandra Rao and P. Ramachandra Rao Pant: Correlation of spread *F* activity with *F*-region height changes

S. I. Akasofu: On the ionospheric heating by hydromagnetic waves connected with geomagnetic micropulsations

V. Marasigan, S. J.: Spread-*F* in Baguio through half a solar cycle

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 181 to 187. Pergamon Press Ltd. Printed in Great Britain

The simultaneous observation of radio star scintillations on different radio-frequencies

H. J. A. Chivers
Jodrell Bank Experimental Station, University of Manchester

(*Received* 2 *May* 1959)

Abstract—A comprehensive study of the phenomenon of radio star scintillations has been made over the range 26 Mc/s to 408 Mc/s, by recording simultaneously on at least two frequencies within this frequency range. With records obtained over a wide range of scintillation conditions, it has been possible to test the theoretical laws governing the frequency dependence of scintillation amplitude and rate. It is established that under extreme conditions these laws no longer hold. A study of the probability amplitude distribution of scintillations covers a wide range of mean scintillation amplitudes. The smaller scintillations have a Gaussian or Rayleigh amplitude distribution, but with increasing scintillation activity, discontinuities appear in the distribution obtained. Under the most intense conditions a nett attenuation of the source is observed. The cross-correlation coefficient between records has been determined over a frequency ratio of 3 : 1 and the results are compatible with those of other workers.

OL.
17
9/60

1. Introduction

There is some disagreement in the results obtained by previous workers from observations on the scintillations of a radio star at different radio-frequencies. Hewish (1952) used frequencies in the range 36 Mc/s to 200 Mc/s and showed that the amplitude of scintillations increases as the square of the observing wavelength whilst scintillation rate is independent of frequency. Bolton *et al.* (1953) also used a large range of frequencies but all their observations were made at angles of elevation below 10°. They found that the scintillation amplitude decreased with increase of frequency up to 100 Mc/s and then appeared to be independent of frequency.

Information on the cross correlation between scintillation records on different frequencies is scanty and contradictory. Burrows and Little (1952) found an average cross-correlation coefficient of 0·61 between records at 118·5 Mc/s and 81·5 Mc/s. They sought for variations in this value with scintillation amplitude or source elevation but found no systematic trend. Other observers have only made general observations on this topic. Bolton and Stanley (1948) found no appreciable correlation between records at 81·5 Mc/s and 100 Mc/s, whilst Smith (1950) occasionally observed good correlation between 45 Mc/s and 81·5 Mc/s records.

In the present paper, a study of scintillations has been made by intercomparison of records at 26 Mc/s, 79 Mc/s, 100 Mc/s and 408 Mc/s. At each frequency, a straightforward total power system was used so that the records obtained were free from interferometer fringes or phase scintillations.

2. Techniques

2.1. 79 Mc/s *and* 26 Mc/s *comparisons*

The 79 Mc/s receiving equipment, normally used for the continuous study of the scintillations of Cassiopeia (24N5A), has been described elsewhere (Dagg,

1957). The aerial is driven about a polar axis so that the source is always in the beam. For the present experiment, an additional output was taken from the receiver so that the routine recordings were continued undisturbed. For the 26 Mc/s observations, the 220 ft diameter transit telescope at Jodrell Bank (Hanbury Brown and Hazard, 1951) has been used in conjunction with a modified commercial communications receiver.

The two equipments used limited the observing programme to a daily observation of Cassiopeia for about 1·5 hr around the time of upper meridian transit of the source. The outputs from the equipments were presented on two channel pen recording milliammeters. On one recorder, the 26 Mc/s scintillations were presented on two sensitivity settings with the chart running at 12 in./hr. On the other recorder the 79 Mc/s and 26 Mc/s signals were presented on strip chart driven at 1 in./min for 5 min of each 15 min period. In addition, the 79 Mc/s scintillations were presented under the routine arrangements on a chart driven continuously at 6 in./hr.

2.2. 100 Mc/s, 79 Mc/s *and* 36 Mc/s *comparisons*

During a recent investigation into the dependence of ionospheric absorption on the elevation of a radio star (Chivers and Greenhow, 1959) records were obtained simultaneously on 100 Mc/s and 36 Mc/s which have now been further analysed to give information on the variation of scintillations at these frequencies. The records have been compared with simultaneous recordings on the routine 79 Mc/s equipment mentioned above (2.1).

VOL.
17
1959/6

2.3. 408 Mc/s *and* 79 Mc/s *comparisons*

The 408 Mc/s survey equipment recently brought into use on the 250 ft radio telescope at Jodrell Bank (Large *et al.*, 1959) was used at low sensitivity on Cassiopeia for about 2 hr to obtain records showing scintillations. Only small scintillations (about 1 per cent of the source intensity) were observed, but these could be measured with high accuracy. The mean values obtained have been compared with the routine recordings on 79 Mc/s.

3. Results

Combined observations on 79 Mc/s and 26 Mc/s were made for seven nights and the results are summarized in Table 1. Upper meridian transit of Cassiopeia during this period was about 2230 UT. The charts driven at constant speed were examined to give the amplitude and rate of scintillations. "Amplitude" is measured as the mean scintillation depth in a $\frac{1}{2}$ hr period expressed as a percentage of the mean source intensity, and "rate" is the number of distinct peaks in the trace per minute. The chart driven at high speed was used to evaluate the cross-correlation coefficient and each 5 min period of running could yield a statistically significant value of the coefficient.

The results of the intercomparison of records on 408 Mc/s, 100 Mc/s, 79 Mc/s and 36 Mc/s are summarized in Table 2. The various aspects of the investigations will now be discussed in some detail.

3.1. *Amplitude of scintillations*

3.1.1. *Wavelength dependence.* Table 1 shows that an index of 100 per cent has been assigned to the 26 Mc/s scintillations for each night of observation. On theoretical grounds (BOOKER, 1958) it can be shown that the scintillation index

Table 1

Date (1958)	26 Mc/s		79 Mc/s		Cross-correlation coefficient
	Fluctuations per min	Amplitude (%)	Fluctuations per min	Amplitude (%)	
30 Sept.	21·0	100	21·0	19	−0·11
1 Oct.	7·5	100	6·8	12	+0·04
2 Oct.	3·0	100	3·0	23	—
3 Oct.	0·7	100	0·7	19	—
4 Oct.	5·2	100	5·2	32	+0·14
5 Oct.	11·0	100	12·0	9	—
6 Oct.	4·8	100	5·2	6	+0·42 +0·33 +0·15

increases as the square of the observing wavelength. Here we have a wavelength ratio of 3 : 1 indicating that the scintillation amplitude on 26 Mc/s should be nine times greater than that on 79 Mc/s.

The smallest scintillation index for 79 Mc/s occurring in Table 1 is only 6 per cent of the mean source intensity. Since, however, the statistical noise ripple on

Table 2

Date (1958)	Frequencies (Mc/s)	Scintillation amplitude (%)	Exponent of wavelength dependence
23 November 0400 UT	408 79	1·2 24·3	1·9
12 April 1750 UT	100 36	6·8 56·0	2·0
12 April 1750 UT	100 79	6·8 10·7	2·1
12 April 1750 UT	79 36	10·7 56·0	2·1

the record has been computed to be 3 per cent of the deflexion due to Cassiopeia, the assignment of any specific scintillation index below about 10 per cent is liable to serious error. On this basis, it is not surprising that whenever small scintillations

are discernible on the 79 Mc/s record, an index of 100 per cent is indicated by the 26 Mc/s record. The theoretical frequency dependence of scintillation amplitude is thus found to hold within the limits of experimental error over the range 26 Mc/s to 79 Mc/s up to the point where the low frequency index saturates at 100 per cent.

Table 2 shows that the frequency dependence of scintillation amplitude also agrees with the theory to within close limits over a range extending from 36 Mc/s to 408 Mc/s. However, the records used to give these results did not exhibit the phenomenon of "saturation" of the scintillation index at 100 per cent. These results appear to disagree with those of BOLTON *et al.* (1953), who found that above 100 Mc/s the scintillation index was independent of frequency. This discrepancy is presumably due to differing observational conditions since the latter results were obtained at low elevations towards the equator, and the present work refers to observations at higher elevations and latitudes.

3.1.2. *Amplitude distribution of scintillations.* A number of scintillation records have been selected to investigate the changes in the probability amplitude distribution of scintillations with scintillation amplitude. The range of activity covered is from 2 to 100 per cent scintillations on 100 Mc/s. For each case considered, the amplitude distribution occurring simultaneously on a lower frequency record has been derived. The results are given in Fig. 1, where the probabilities are plotted as a function of fractional source intensity.

Considering the higher frequency first, Fig. 1 (a), (c), (e), (g) and (j), it is apparent that the amplitude distribution for small scintillations is a Gaussian curve about the mean source intensity. The width of this curve increases with the scintillation activity, until for very deep scintillations, Fig. 1 (j), the resultant is no longer Gaussian in shape but appears to be a Rayleigh distribution. This last curve has been tested using a method given by FÜRTH and MCDONALD (1947). RICE (1944, 1945) has shown that for a statistically random phenomenon the probability $p(R)\,dR$ for values of amplitude R within the interval $(R,\ R+dR)$ is given by

$$p(R) = \frac{R}{\psi}\exp(-R^2/2\psi)$$

A sensitive test for a distribution of this type is to plot experimental values of $\ln p(R)/R$ against R^2. This should yield a straight line of slope $-1/2\psi$ if the theoretical formula is obeyed. Using the value of ψ so obtained, a theoretical curve of $p(R)$ against R can be compared with the experimental curve. From Fig. 1 (j) it can be seen that the deep 100 Mc/s scintillations have an amplitude distribution which follows closely the theoretical curve for a Rayleigh distribution.

Examining the amplitude distributions of the lower frequency records, Fig. 1 (b), (d), (f), (h), (k), it is seen that, again, the smaller scintillations have a Gaussian amplitude distribution. For larger amplitudes a skew distribution is obtained, and in each case a theoretical curve derived as above has been superimposed on the experimental curve. Fig. 1 (d) shows that in the presence of 9 per cent 79 Mc/s scintillations, the 26 Mc/s records do not exhibit a fully developed Rayleigh distribution, but one that is nearly Gaussian. For more intense scintillation activity, the curves become irregular in shape and do not follow closely a theoretical Rayleigh distribution. Fig. 1 (k) shows that when scintillation activity is such that an index

VOL
17
1959/

of 100 per cent is assigned to the 100 Mc/s record, the mean signal level on 36 Mc/s is only a fraction of the expected source intensity.

The range of scintillation activity demonstrated in Fig. 1 can best be described in terms of the scattering efficiency of the diffracting screen. When weak scattering occurs, the mean amplitude of the scattered waves is small compared with the unscattered remnant of the incident wave. The resultant amplitude observed at the ground consists of the unscattered wave plus the sum of the scattered waves from

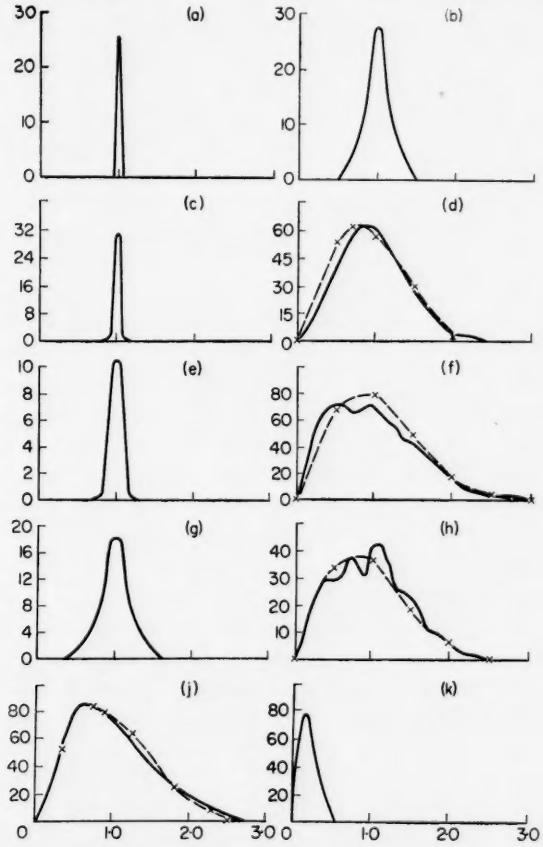


Fig. 1. Probability amplitude distribution of scintillations. ———— Experimental curves, — — — — theoretical curves. Ordinates, probability $p(R)$, (arbitrary units). Abscissae, intensity R, (fraction of source intensity).

(a) 12 April 1615 UT 100 Mc/s. Mean amplitude 2 per cent.
(b) 12 April 1615 UT 36 Mc/s. Mean amplitude 20 per cent.
(c) 5 Oct. 2230 UT 79 Mc/s. Mean amplitude 9 per cent.
(d) 5 Oct. 2230 UT 26 Mc/s. Mean amplitude 100 per cent.
(e) 2 Oct. 2215 UT 79 Mc/s. Mean amplitude 23 per cent.
(f) 2 Oct. 2215 UT 26 Mc/s. Mean amplitude 100 per cent.
(g) 11 April 2050 UT 100 Mc/s. Mean amplitude 30 per cent.
(h) 11 April 2050 UT 36 Mc/s. Mean amplitude 100 per cent.
(j) 12 April 1805 UT 100 Mc/s. Mean amplitude 100 per cent.
(k) 12 April 1805 UT 36 Mc/s. Mean amplitude ?

the irregularities in the cone of reception. Since the scattered waves have a random phase distribution with respect to the unscattered wave, the resultant will be a Gaussian amplitude distribution about the mean source amplitude, e.g. Fig. 1 (a), (b), (c), (e), (g). Under conditions where the mean amplitude of the scattered waves becomes comparable with the mean source intensity, the statistical sum of the scattered radiation will have a Rayleigh distribution about the mean source intensity (Fig. 1 j). If the ionosphere is particularly disturbed, some of the irregularities in the cone of reception behave as electron lenses at the lower frequencies considered. These cause focusing and defocusing of the incident wave which will give discontinuities in the probability amplitude distribution because there will be insufficient lenses in the cone of reception to give a statistically random amplitude distribution (Fig. 1 f, h). Should conditions be such that strong scintillations are observed on a frequency as high as 100 Mc/s, it is clear from Fig. 1 (k) that the irregularities in the ionosphere are sufficient to cause a nett attenuation of the signal at 36 Mc/s. These surprising conditions appear to be common when observing sources at low elevations through the auroral zone (CHIVERS and GREENHOW, 1959).

3.2. *Rate of scintillations*

The results of the 26 Mc/s and 79 Mc/s experiment given in Table 1 show that scintillation rates are very similar even though a large day to day variation in scintillation was observed. However, in comparing rates for the observations in Fig. 1, it has been found that the observed rates are the same on two frequencies only when scintillation activity is slight (Fig. 1 a, b, c, d, e, f). When 30 per cent scintillations are observed on 100 Mc/s (Fig. 1 g), the rate on 36 Mc/s is found to be about double that on 100 Mc/s. For the more intense conditions leading to Fig. 1 (j), (k), a comparison is not possible because of the severe attenuation on the lower frequency.

It has been pointed out by HEWISH (1951) that if the random phase changes introduced by an irregular diffracting screen are less than 1 rad, the scale of the diffraction pattern produced at the ground has the same lateral scale as the irregularities in the screen, so that the observed scintillation rate will be the same on any frequency for which this condition holds. However, when phase changes in excess of 1 rad occur, a fine structure is produced in the diffraction pattern having a scale smaller than that in the screen. This will cause an increase in scintillation rate with observing wavelength since the change in scale will become finer with increasing wavelength. It is apparent that when ionospheric irregularities are capable of producing 30 per cent scintillations on 100 Mc/s (Fig. 1 g), then the phase variations in the 36 Mc/s wave exceed 1 rad, leading to an increased scintillation rate at the lower frequency.

3.3. *Cross-correlation analysis*

The values of the cross-correlation coefficient which have been evaluated for a few of the 79 Mc/s and 26 Mc/s records are exceedingly low and below the level of a statistically significant result. This evaluation can only be compared with the work of BURROWS and LITTLE who found a mean coefficient of 0·61 using a frequency ratio of 1·5 : 1.

VOL.
17
1959/6

Linear extrapolation from the latter result indicates that an insignificant correlation coefficient should be obtained for a frequency separation of about 2·5 : 1. The present result does not add any more data on the rate of change of the correlation coefficient with frequency separation, but at least establishes that with a frequency ratio of 3 : 1 the coefficient has dropped to an insignificantly low value.

4. Conclusions

Various aspects of the phenomenon of the scintillation of a radio star have been discussed in the light of experiments covering the range 26 Mc/s to 408 Mc/s. It has been shown that:

(i) The mean scintillation amplitude varies as the square of the observing wavelength up to the point where the lower frequency index saturates at 100 per cent. For zenithal observations on 26 Mc/s at the latitude of Jodrell Bank (53° N), an index of 100 per cent is commonly obtained.

(ii) The rate of scintillations is independent of observing wavelength under weak scattering conditions, and for strong scattering, the rate increases with wavelength. Both of these effects are predicted theoretically.

(iii) The probability amplitude distribution of small scintillations is a displaced Gaussian curve, and larger amplitudes have a Rayleigh distribution. Very intense scintillations have an irregular amplitude distribution indicating the occurrence of non-statistical focusing and defocusing of the incident wave.

(iv) There is no significant cross-correlation between records of scintillations obtained simultaneously on frequencies differing in the ratio 3 : 1. Combining this with the results of other workers, it is clear that the fall in cross correlation with frequency separation is at least at a linear rate.

Acknowledgements—The writer is indebted to Professor A. C. B. Lovell who encouraged the production of this paper, and to Dr. J. S. Greenhow for his valuable advice. Thanks are due to many colleagues who made the recordings on various frequencies possible.

References

Bolton J. G. and Stanley G. J.	1948	*Aust. J. Sci. Res.* A **1**, 58.
Bolton J. G., Slee O. B. and Stanley G. J.	1953	*Aust. J. Phys.* **6**, 434.
Booker H. G.	1958	*Proc. Inst. Radio Engrs, N.Y.* **46**, 298.
Burrows K. and Little C. G.	1952	*Jodrell Bank Annals* **1**, 29.
Chivers H. J. A. and Greenhow J. S.	1959	*J. Atmosph. Terr. Phys.* **17**, 1.
Dagg M.	1957	*J. Atmosph. Terr. Phys.* **10**, 204.
Fürth R. and MacDonald D. K. C.	1947	*Proc. Phys. Soc.* **59**, 388.
Hanbury Brown R. and Hazard C.	1951	*Mon. Not. R. Astr. Soc.* **111**, 357.
Hewish A.	1951	*Proc. Roy. Soc.* A **209**, 81.
	1952	*Proc. Roy. Soc.* A **214**, 494.
Large M. I., Matthewson D. and Haslam G.	1959	*Nature, Lond.* **183**, 1250.
Rice S. O.	1944	*Bell Syst. Tech. J.* **23**, 282.
	1945	*Bell Syst. Tech. J.* **25**, 46.
Smith F. G.	1950	*Nature, Lond.* **165**, 423.

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 188 to 204. Pergamon Press Ltd. Printed in Great Britain

The azimuth distribution of oblique reflections from meteor trails and its relation to meteor radiant distributions

W. C. Bain
D.S.I.R. Radio Research Station, Slough*

(*Received* 5 *May* 1959)

Abstract—Azimuth distributions of the meteor reflections to be expected at various times of day on a 1740 km north–south transmission path at v.h.f. are calculated from three initial radiant distributions of sporadic meteors. These distributions were: (1) a distribution of meteors, magnitude about 7, obtained by workers at Jodrell Bank; (2) a uniform heliocentric distribution in the plane of the ecliptic; (3) a heliocentric distribution uniform over the celestial sphere. It was found that the third of these gave the best fit to the azimuth results obtained by Meadows (1958) on such a path at a frequency of 37 Mc/s; the meteors observed extended downwards in intensity to magnitude 10. Calculations on the forward-scattering of radio waves by meteors should therefore be based on this distribution.

Observational results are also presented for a very similar path with a frequency of about 70 Mc/s; they resemble those obtained at 37 Mc/s, although are not identical to them. The cause of the differences is unknown, but they may be due to the transmitter polar diagram not being identical at the two frequencies.

1. Introduction

Results on the direction of arrival of reflections from meteor trails at v.h.f. over a long transmission path have already been given by Meadows (1958). In that work the transmitter used had a frequency of about 37 Mc/s and was situated at Gibraltar, the receiving site being at Winkfield, near Windsor. Subsequently the frequency of this transmitter was changed to a value near 70 Mc/s, and the directional measurements were continued on a site at Slough about 10 km from the original one. The bearing of Gibraltar from this point was 194° and its distance 1740 km. Section 2 of this paper contains an account of the azimuth results obtained on the higher frequency.

The connexion between the azimuth distribution of the reflections and the distribution of sporadic radiants over the celestial sphere is of considerable interest. If the radiant distribution were known, then the azimuth distribution of meteor reflections which would be obtained on any transmission path on the surface of the earth could be calculated, and this would enable one to design aerials for a system of the JANET type, for instance. Unfortunately the distribution of radiants is not accurately known, as is pointed out by Forsyth *et al.* (1957). The method adopted here has therefore been to calculate the expected azimuth distribution from three different possible meteor radiant distributions, and to compare these results with the actual observed distributions on 37 and 70 Mc/s. The distributions chosen were:

(1) A distribution derived by the workers at Jodrell Bank Experimental Station from their radar observations on meteors of low magnitude.
(2) A heliocentric distribution uniform over the ecliptic plane but confined to that plane.

* Official communication.

VOL.
17
1959/6

(3) A heliocentric distribution uniform over the whole celestial sphere.

Considerable approximations were introduced into the analysis to simplify the calculations. This appears reasonable, as the meteor frequencies with which the results were to be compared cannot be supposed to be established to an accuracy greater than 20 per cent.

2. Azimuth Observations on 70 Mc/s

2.1. *Equipment*

The transmitter at Gibraltar on 70 Mc/s had a nominal power of 30 kW and its main aerial had theoretically a 3 dB beamwidth of about 24° centred on a bearing of 12° east of north. It should be noted, however, that the transmitter aerial for 37 Mc/s was designed to have these properties, but was found by measurement to have its main lobe displaced westwards by about 7°; hence too much reliance cannot be placed on the theoretical figures. This aerial was at a height of 600 ft above sea level so that the lobes of the polar diagram produced by earth reflection would be spaced by about 0·6° in elevation. These lobes are so close together that the azimuth distribution of meteor reflections at Slough will be practically the same as for a uniform elevation polar diagram. Further details about the transmitter are given by Williams (1958).

The receiving aerial system and equipment employed for this work were set up primarily to receive the ionospheric scatter signal from Gibraltar on 70 Mc/s, and were very similar (in parts identical) to that used on 37 Mc/s and described in an earlier paper (Bain, 1958). The signals were picked up at Slough on two horizontal half-wave dipoles at a height of 23 m at right angles to the great circle path from Gibraltar, taken to balanced receivers, and from there to phase-measuring equipment of the sum-and-difference type. A narrow bandwidth of 25 c/s was used in the receivers to enable the meteor selection level to be fixed at the low value of -160 dB W approximately. This was necessary because of the much weaker signals received on 70 Mc/s compared with 37 Mc/s.

For the actual observations of meteors the aerial spacing was set to one wavelength and the phase-and-amplitude balance of the receivers was adjusted with the aid of radiation from a local symmetrically placed aerial. The great majority of the results presented here were obtained by visual reading of the scale on the cathode-ray tube showing the phase display. The gain of the receivers was set to bring a signal at the level of -160 dB W to a given size on the screen, and any signal from Gibraltar exceeding this level was taken to be due to a meteor. On the sum-and-difference type of display the signal appears as an ellipse whose major axis makes an angle of $\phi/2$ with a fixed direction, ϕ being the phase difference between the signals reaching the two aerials. This angle was read once, as quickly as possible after the appearance of each meteor, with the aid of a rotating cursor. Most meteor signals faded out before the reading could be taken, so a screen with a few seconds' afterglow was employed; this enabled the angle to be read quite easily. A small number of periods of photographic recording was also carried out.

Each reading of $\phi/2$ taken from the c.r.t. was converted to the bearing α by the formula

$$\phi = 360 \sin \alpha \quad (\phi \text{ in degrees})$$

as the aerial spacing was one wavelength and the angle of elevation was not expected to exceed about 8°.

2.2. *Results*

Observations of the azimuths of meteor trails were made from March to May 1957; this time of year is the same as that during which the main part of the 37 Mc/s results were taken in 1956. Histograms were constructed for 3 hr groups, and in each the number of observed bearings contained in blocks 3° wide was plotted against the

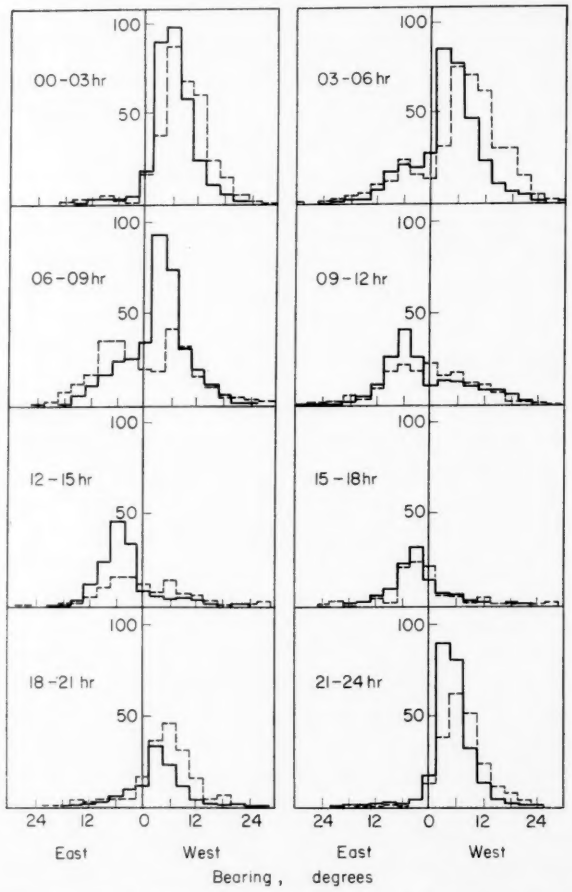


Fig. 1. Bearings of meteor reflections on 70 and 37 Mc/s. Rate per hour in each 3° bearing group against bearing relative to the great-circle direction. ——— 70 Mc/s; – – – 37 Mc/s.

bearing, as shown in Fig. 1. The bearing shown as zero corresponds to the great-circle direction of the transmitter. For comparison, the 37 Mc/s results here are shown by dotted lines; their rate/hr has been adjusted to the same average value as that for 70 Mc/s.

The results on the two frequencies are quite similar. There is a strong tendency in both for the bearings to be westward at night and eastward during the day,

VOL. 17 1959/6

with a minimum of activity occurring near 1800 UT. The main differences lie in the higher values at bearing 3° in the periods from 2100 to 0900 hours for 70 Mc/s, which are rather difficult to account for, and in the somewhat sharper cut-off at the higher westerly bearings which is evident in the 70 Mc/s results. This latter feature might be due to an unintentional change in the transmitter aerial polar diagram between the two frequencies; it should be noted that the main lobe in the 37 Mc/s polar diagram is about 7° farther to the west than the theoretical direction of the 70 Mc/s lobe.

A check on the performance of the direction-finding equipment was carried out occasionally by setting the aerial spacing to two wavelengths instead of one. The

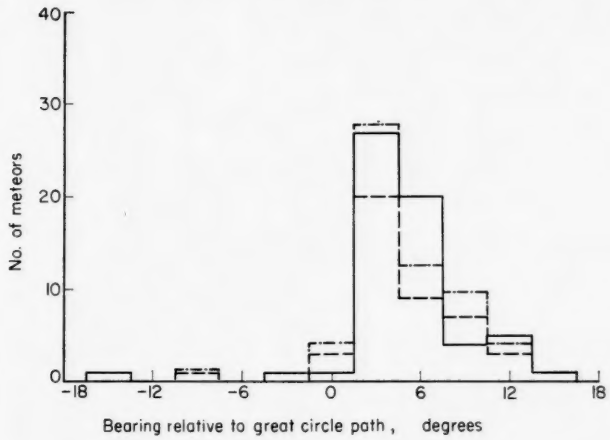


Fig. 2. Azimuth distributions obtained in two adjacent 10 min periods near 2200 hours UT on 15 April 1957. – – – 1λ spacing; ——— 2λ spacing; –·–·– 1λ spacing (average rate adjusted to the same as for 2λ).

azimuth distribution was calculated for this period and compared with that for an equal adjacent period of time when the one-wavelength spacing was used. The agreement between the two was satisfactory; a particular example is shown in Fig. 2.

3. The Calculation of Azimuth Distributions

3.1. *General*

A calculation has already been made by Hines (1955) of the number of meteors counted per unit time for a fixed meteor radiant position, and the validity of his approximations for transmission paths of over 1000 km in length has been confirmed in another paper by Hines and Pugh (1956). The expression given by Hines is

$$F_0 = \frac{\sin^2\beta \cot\psi\,[(\mathscr{D}^2 - h^2/\sin^2\psi)^{\frac{1}{2}} - D]\,(D^2 - h^2\cot^2\psi)}{r^{\frac{3}{2}}[1 + h/\{(R+H)\sin^2\psi\}]\,(1-\cos^2\beta D^2/r^2)^{\frac{1}{2}}\,(D^2 + h^2\cot^2\psi)} \qquad (1)$$

Here F_0 is a number proportional to the number of meteors observed above a certain level;

β, ψ are co-ordinates defining the position of the meteor radiant relative to the transmission path;
R is the radius of the earth;
H is the height of the meteor trail (taken here as 100 km);
$\mathscr{D} = [(R+H)^2 - R^2]^{\frac{1}{2}}$;
$2D$ is the straight-line chord distance from transmitter to receiver;
h is the height of the trail above a plane through transmitter and receiver parallel to the horizontal at their midpoint, and is a function of ψ but not of β.

Now all the reflection points due to meteors with the radiant ψ, β which reflect signals to the receiver must project on to the above plane through TR (see Fig. 3) into points lying on a line AB which is parallel to TR and a distance h cot ψ from

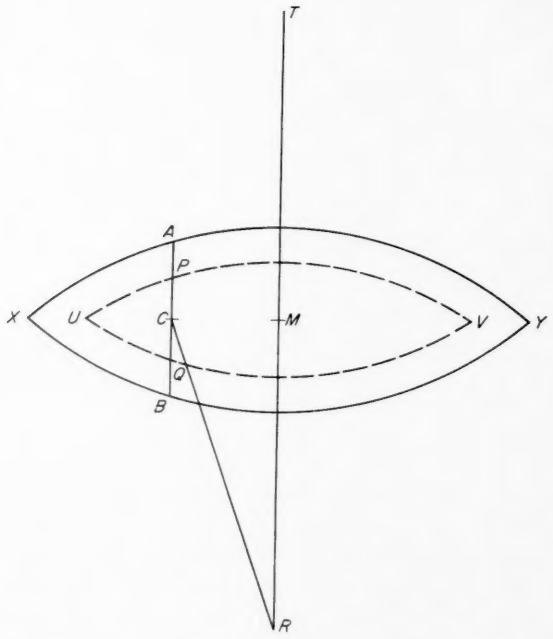


Fig. 3. Projection on the plane through the transmitter (T) and receiver (R) parallel to the horizontal at their mid-point.

it. The situation is illustrated by Fig. 3, in which the area contained by the arcs joining X to Y is a projection of that area at 100 km height on which each point is at an angle of elevation greater than 0° from both T and R. To take into account the polar diagram of the transmitting and receiving aerials Hines integrated an illumination factor G along this line. In the present case the variation in the receiver polar diagram with azimuth may be taken as negligible. The transmitter polar diagram has been measured with the aid of an aircraft (Crow *et al.*, 1956), and for use here its main lobe has been represented by a sinusoidal function with its maximum directed 5° to the west of the Slough bearing and with zeros on bearings 25° on either side of this direction. The best polar diagram in elevation

VOL. 17 1959/6

to use for these aerials is more difficult to decide. However, the results obtained by MEADOWS show that very few reflections were obtained at angles of elevation less than 1°, and the polar diagram of the receiving aerial in elevation has accordingly been taken to be zero below this angle but uniform above it. The same assumption has been made for the transmitter aerial as in that case the radiation was cut off at about 1° elevation by a range of hills. This reduces the projection of the common area at 100 km height visible from both transmitter and receiver to that contained by the dotted lines between U and V in Fig. 3; hence only the length PQ of the line AB will be effective in giving reflections. Allowing for these polar diagram factors, the meteor count rate becomes $F_0{}^*$ where

$$F_0{}^* = F_0 \int G\,dx \tag{2}$$

the integration being along PQ. To simplify the calculation the value of G at any point on PQ has been taken as equal to its value at the mid-point C of PQ. The errors introduced by this should not be large as the length PQ is always much shorter than TR. The integral therefore becomes equal to $G \times$ length PQ.

For a reflection point equidistant from T and R the bearing of the incident wave relative to the transmitter bearing is given by γ where

$$\tan\gamma = \frac{-h\cot\psi}{D\cos Q + h\sin Q} \tag{3}$$

Q is half the angle subtended by TR at the centre of the earth.

The negative sign arises because the reflection is always on the opposite side of TR to the direction of the radiant. Now to an accuracy everywhere better than $\frac{1}{2}°$,

$$\tan\gamma = -\frac{h\cot\psi}{D} \tag{4}$$

i.e. the projected angle (CRM in Fig. 3) may be taken as the bearing. In the subsequent analysis the bearings have been divided into 3° groups. Clearly PQ may subtend more than 3° at R if $h\cot\psi$ is not small, and the bearings arising from a fixed value of ψ have accordingly been allocated to more than one of these groups where necessary.

For a long-distance transmission path the great majority of the reflections must be ascribed to sporadic meteors, whose radiants do not coincide, but are spread over the celestial sphere. This radiant distribution is most conveniently expressed as a function of ecliptic latitude and longitude relative to the apex of the earth's way; it will be assumed in this paper to be independent of time.

To calculate bearings, the values of ecliptic latitude and longitude relative to the apex must be related to the co-ordinates β and ψ. This can be done in stages as follows. For a particular time of year, these ecliptic co-ordinates can be converted to right ascension (α) and declination (δ). For any given time of day the right ascension can be converted to hour angle (H) at the path mid-point, and the azimuth (A) and zenith angle (z) can then be found for this position. The correction for the gravitational attraction of the earth has not been applied, as it

is not likely to increase significantly the accuracy of the results. The azimuth A relative to true north can then be converted to azimuth θ relative to the TR direction by

$$\theta = A - A_0$$

where A_0 is the bearing of Gibraltar from the path mid-point. The co-ordinates β, ψ can now be obtained from θ, z by the relations

$$\tan \psi = \sin \theta \tan z$$

$$\cos \beta = \sin z \cos \theta$$

The meteor count rate F_0* and the associated bearings γ can then be obtained from equations (2) and (4).

In practice contour charts of F_0 and of γ for the Gibraltar–Winkfield path have been drawn as functions of H and δ; the chart for F_0 is shown in Fig. 4. Both charts are in polar form, with $(90 - \delta)$ as the radial co-ordinate; they can be associated

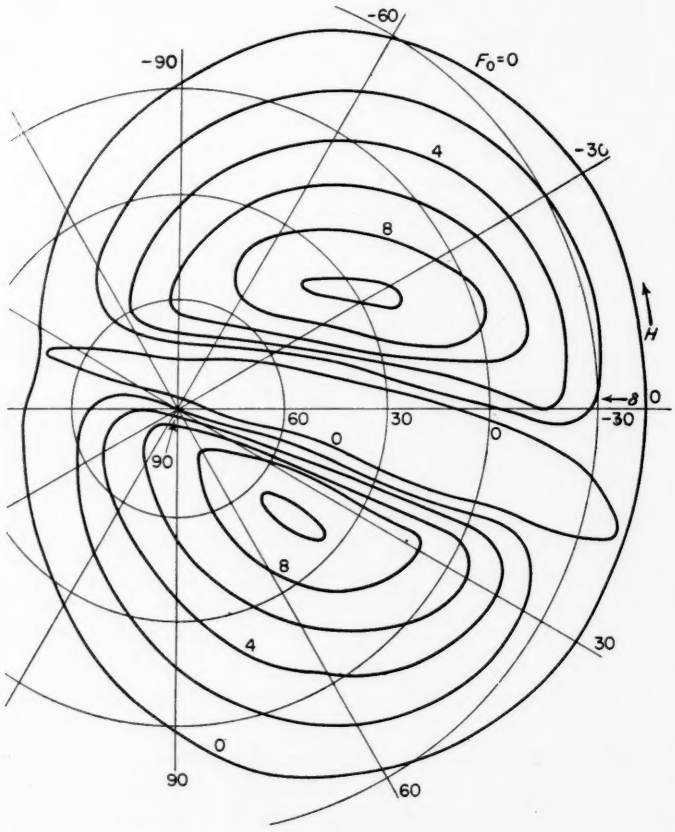


Fig. 4. Contour chart for F_0 at intervals of 2, for values of hour angle (H) and declinations (δ) in degrees.

with meteor distributions as follows. Each meteor distribution given in ecliptic latitude and longitude has been simplified to a finite number of line-distributions in these co-ordinates and a particular weight has been associated with each line. The lines have then been converted to corresponding lines in right ascension and declination, and have been drawn on tracing paper to the same scale as the F_0 and γ charts, but with right ascension in place of hour angle. As the sum of the right ascension and the hour angle is the local sidereal time, the tracing can be superimposed on either the F_0 or γ charts and adjusted to the position corresponding to any particular time of day by a simple rotation. The count rates and bearings can then be plotted out on a separate sheet. The rates are then multiplied by the factor $\int G\,dx$ and by the weight, and a consolidated azimuth distribution for that time of day obtained; this is then finally corrected by allowing for the spread in bearing for fixed ψ.

In the analysis of the observations, the distributions have been taken over 3 hr periods. In the calculations, therefore, the distributions have usually been found for the times 0000, 0300, 0600, ..., etc., hours UT. The distribution during each period has been taken to be proportional to the mean of the distribution for each end of that period.

3.2. *Actual distributions used*

The distribution of apparent meteor radiants produced by the Jodrell Bank workers (Lovell, 1954; Davies, 1957) is shown in Fig. 5. It contains two main lobes lying in and near to the ecliptic plane (curve R), with, in addition, an appreciable proportion of radiants out of the ecliptic plane in the direction nearest to the apex of the earth's way (the difference between curves P and R). A similar distribution was used by Kent (1958) in his work on the Kirk o' Shotts to Cambridge path; he found that it agreed well with the results. Here this distribution has been simplified to a series of lines on the celestial sphere, each associated with a weight proportional to the number of radiants as shown in Table 1.

Table 1. Distribution A (simulating the Jodrell Bank distribution)

Celestial latitude (deg)	Longitude from apex (deg)		Weight
	From	To	
0	280	310	1·00
0	330	20	0·43
0	45	80	0·58
30	330	30	0·43
60	330	30	0·26

The tabulated weight is the total weight of that particular line-distribution.

The next distribution considered was the uniform heliocentric distribution confined to the plane of the ecliptic. The parabolic heliocentric velocity for a meteor at the earth's distance from the sun is 42·1 km/sec. Various values have

been taken for a typical heliocentric velocity for a meteor, the parabolic velocity itself being frequently chosen. As modern work indicates clearly that most meteors have in fact less than this velocity, the value of 40 km/sec has been used here for

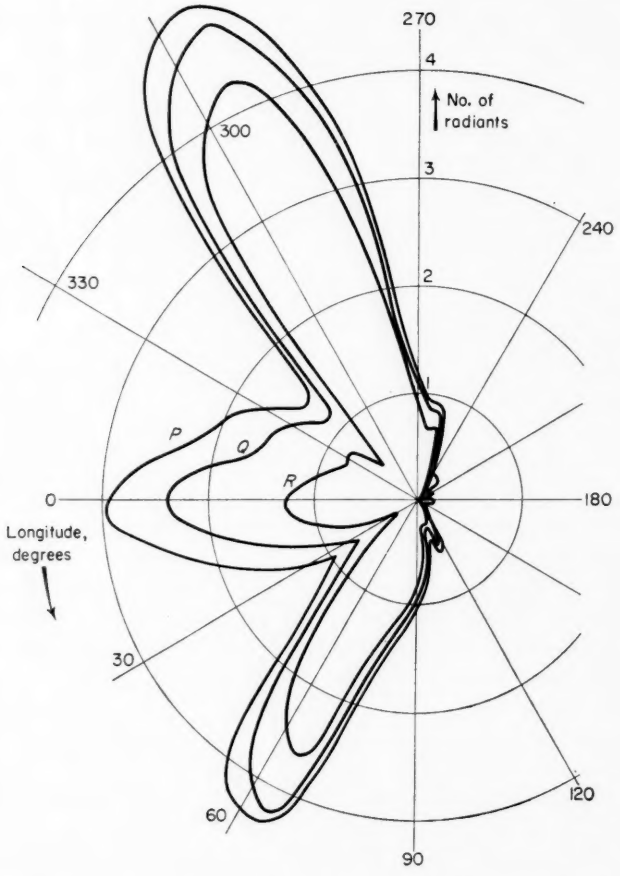


Fig. 5. The number of sporadic meteor radiants (mainly of magnitude 7 and 8) as a function of celestial longitude relative to the apex, obtained by the Jodrell Bank workers.

convenience, the earth's velocity being taken as 30 km/sec. The meteor radiant distribution then becomes

$$F_1(\theta)\,d\theta = \frac{1}{2\pi}\left[1+\frac{\cos\theta}{\sqrt{(0{\cdot}78+\cos^2\theta)}}\right]d\theta \qquad (5)$$

θ is an angle in the ecliptic plane measured from the apex, and $F_1(\theta)$ is the normalized distribution function. It is shown in Fig. 6 together with the line-distributions forming distribution B, which is an approximation to $F_1(\theta)$ and has been used in the calculations. The details of distribution B are given in Table 2.

VOL.
17
1959/6

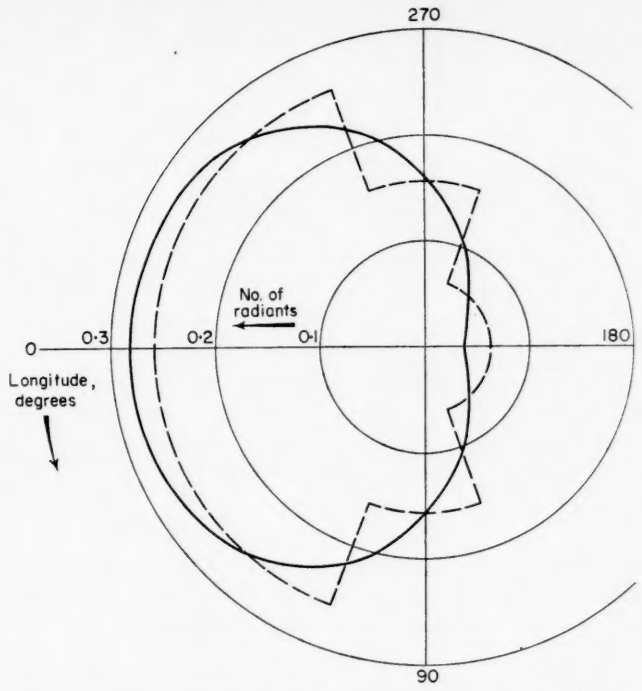


Fig. 6. The number of meteor radiants as a function of celestial longitude relative to the apex for a heliocentric distribution confined to the ecliptic. ——— F_1; – – – distribution B, simulating it.

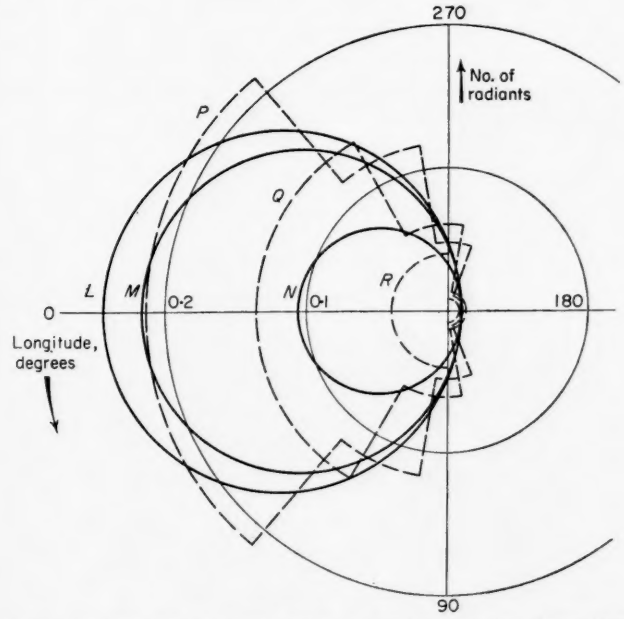


Fig. 7. The number of meteor radiants as a function of celestial longitude relative to the apex (α), for a heliocentric distribution over the whole celestial sphere. P, Q and R simulate meteors in the ranges of β' from -20 to 20, 20 to 50, and 50 to 90 degrees.

Table 2. Distribution B (simulating a uniform heliocentric distribution confined to the ecliptic)

Ecliptic latitude (deg)	Longitude from apex (deg)		Weight
	From	To	
0	290	70	1·00
0	70	110	0·18
0	110	250	0·25
0	250	290	0·18

The tabulated weight is the total weight of the line-distribution.

For the uniform heliocentric distribution over the whole celestial sphere the same value was used for the meteor velocity. The meteor distribution then becomes

$$F_2(\theta)\,d\theta = \tfrac{3}{8}\sin\theta\sqrt{(0{\cdot}78+\cos^2\theta)}\left[1+\frac{\cos\theta}{\sqrt{(0{\cdot}78+\cos^2\theta)}}\right]^2 d\theta \qquad (6)$$

where θ is now the angle from the apex to the radiant. This can be converted into the more convenient co-ordinates of ecliptic latitude (β') and longitude relative to the apex (a). The meteor distribution $F_3(a,\,\beta')$ becomes

$$F_3(a,\beta')\,da\,d\beta' = (3/16\pi)\sqrt{(0{\cdot}78+\cos^2 a\cos^2\beta')} \times\left[1+\frac{\cos a\cos\beta'}{\sqrt{(0{\cdot}78+\cos^2 a\cos^2\beta')}}\right]^2 \frac{\sin^2 a+\cos^2 a\sin^2\beta'}{\cos\beta'\,(\sin^2 a+\tan^2\beta')}\,da\,d\beta' \qquad (7)$$

This distribution is illustrated by Fig. 7, which also shows the approximating line-distributions C as listed in the following table.

Table 3. Distribution C (simulating a uniform heliocentric distribution over the whole celestial sphere)

Ecliptic latitude (deg)	Longitude from apex (deg)		Weight
	From	To	
0	310	50	1·00
0	50	80	0·17
0	80	110	0·07
0	110	250	0·08
0	250	280	0·07
0	280	310	0·17
35	300	60	0·58
35	60	100	0·08
35	100	260	0·08
35	260	300	0·08
65	270	90	0·32
65	90	270	0·07

The tabulated weight is the total weight of the line-distribution.

VOL. 17 1959/6

4. Comparison of Actual and Calculated Azimuth Distributions

The polar diagram of the transmitter aerial on 70 Mc/s has never actually been measured, so it seems that the comparison between actual bearing distributions and those calculated can best be made for the 37 Mc/s frequency. These results are therefore mainly used in what follows, although some account is given of the agreement with the 70 Mc/s results.

Figs. 8, 9 and 10 show the actual results for 37 Mc/s compared with the bearing distributions calculated from the radiant distributions A, B and C, respectively.

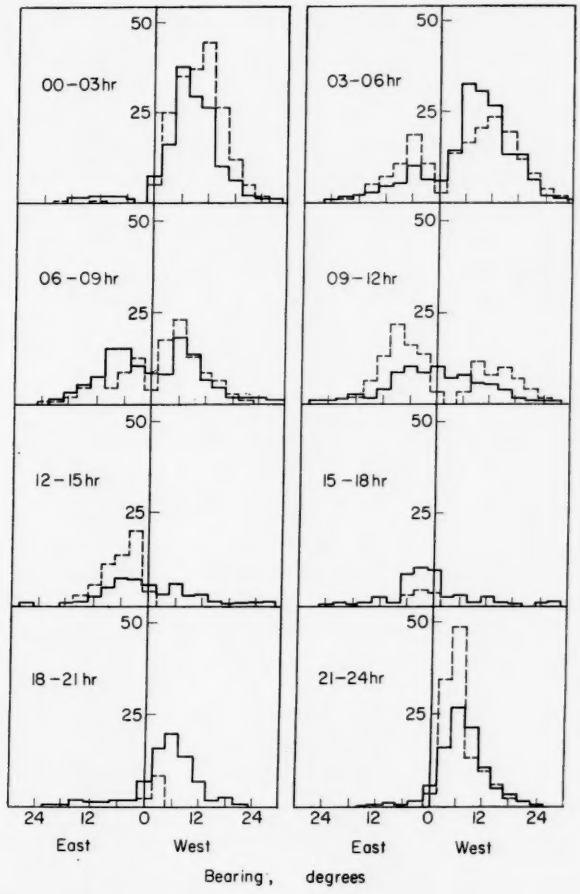


Fig. 8. Comparison of the observed azimuth distribution with that calculated from radiant distribution A, which simulates the Jodrell Bank distribution. ——— observed on 37 Mc/s; – – – calculated.

The calculated distributions have each been multiplied by a constant to bring them to the same total rate over a day as the actual observations. It should also be mentioned here that, in the case of distribution A, the change between the distributions for times separated by 3 hr was found to be too great for the mean

of these two distributions to represent well that for the whole 3 hr period. This effect is caused by the very limited range of longitude occupied by the most important line-distributions in A. Hence for this radiant distribution only, the azimuth distribution at the mid-point of each 3 hr period was calculated and used to improve the approximation to the distribution for the whole period.

By inspection it can be seen that distribution C provides the best fit to the azimuth results. However, it is desirable to confirm this by some test. Now none

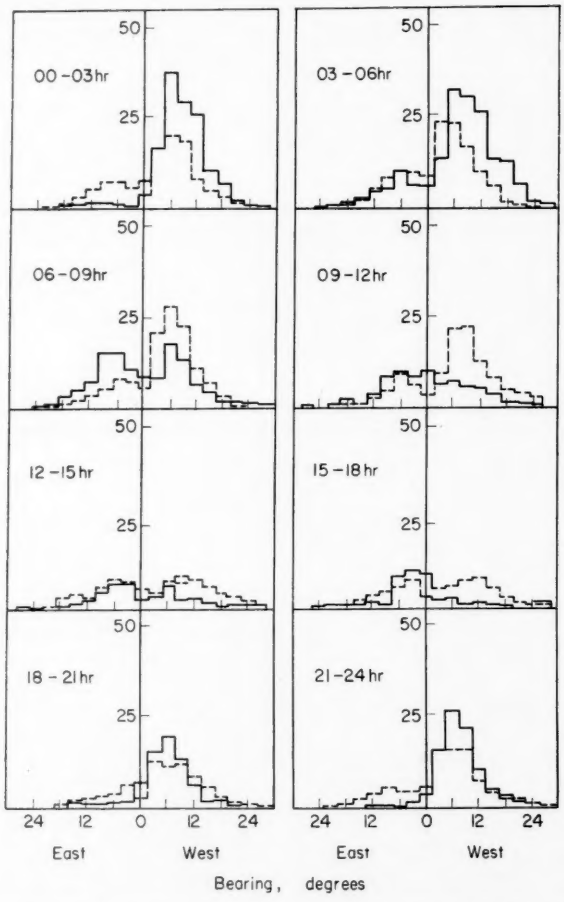


Fig. 9. Comparison of the observed azimuth distribution with that calculated from radiant distribution B, which simulates the uniform heliocentric distribution in the ecliptic plane. ——— observed on 37 Mc/s; – – – calculated.

of the distributions is a sufficiently good fit to warrant the use of a χ^2 test with its accompanying significance levels. Instead the quantity w^2 (Kendall, 1948) has been calculated for each distribution, where

$$w^2 = \sum_i (f_i - m_i)^2$$

VOL. 17 1959/6

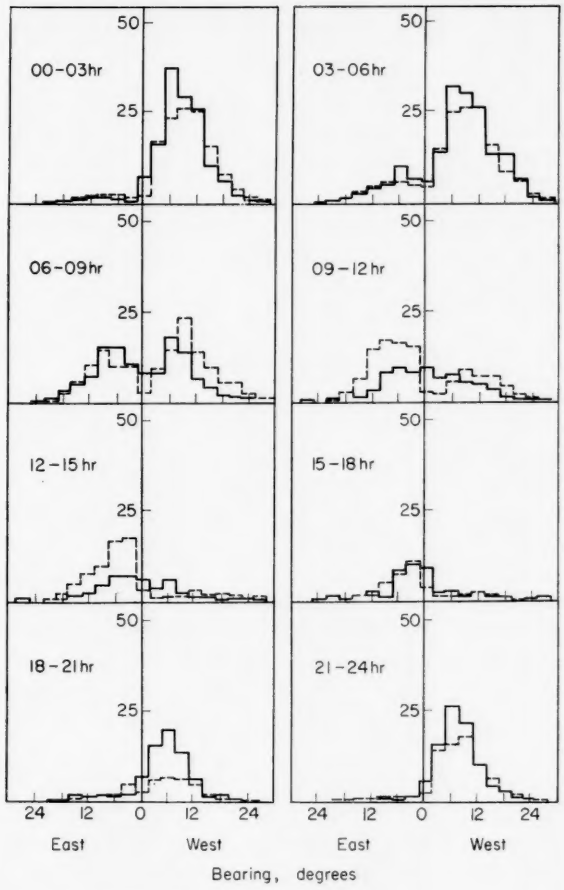


Fig. 10. Comparison of the observed azimuth distribution with that calculated from radiant distribution C, which simulates a heliocentric distribution uniform over the celestial sphere. ——— observed on 37 Mc/s; – – – calculated.

f_i is the actual meteor frequency for a particular 3 hr period and 3° range of bearing; m_i is the corresponding predicted meteor frequency.

The summation is taken over all periods in the day and over all bearing ranges where non-zero predicted frequencies occur at some time. Clearly the smaller is w^2, the better is the fit. The values actually obtained are shown in Table 4.

This confirms that distribution C gives easily the best fit; assuming that the distributions A, B and C are reasonable representations of the original ones from which they were approximated, it follows that the uniform heliocentric distribution of radiants gives considerably better results than the uniform distribution confined to the ecliptic or the distribution from Jodrell Bank. Of the latter two, the distribution confined to the ecliptic will not be considered further. However, reasons must be sought for the poor fit of the Jodrell Bank distribution, which is based on

Table 4. Goodness of fit of azimuth distributions A, B and C to the 37 Mc/s results

Values of w^2 (hr^{-2})

Dist. \ Time (UT)	0000–0300	0300–0600	0600–0900	0900–1200	1200–1500	1500–1800	1800–2100	2100–2400	Totals
A	781	566	333	598	342	137	744	897	4400
B	925	1005	605	745	223	248	188	240	4180
C	277	148	292	443	301	70	353	156	2040

experimental results. In the first place it appears to be partly due to absence of radiants in it with longitudes in the 180° sector opposite to the apex. For instance, this causes the poor fit in the 1800–2100 hours period (see Fig. 8 and the appropriate column of Table 4). It may be noted here, too, that WEISS (1957) found at least a component of sporadic meteors uniformly distributed over the celestial sphere.

One reason for the lack of fit with the Jodrell Bank results might be that the meteors giving the reflections on the Gibraltar–Winkfield path are of a different magnitude to those studied at Jodrell Bank, which were mainly of magnitude about 7. To determine the magnitude of the meteors observed by forward scatter a calculation similar to that given by MEADOWS (1958) may be used. From a knowledge of the received power of the weakest meteors observed, it is found that the meteor-trail reflection coefficient (g) for these meteors is 0·052. This is related to the electron line-density (q) in the trail by the formula

$$g = 8{\cdot}85 \times 10^{-15} q \frac{\sin \alpha}{(1 - \cos^2 \beta \sin^2 \phi)^{\frac{1}{2}}}$$

where α, β, ϕ are geometrical parameters defined by MEADOWS. The expression $\sin \alpha / (1 - \cos^2 \beta \sin^2 \phi)^{\frac{1}{2}}$ varies from 1·5 to 5·5, but the higher values will be favoured as the reflection coefficient is proportional to it. Taking its value as 5·5, this gives the electron line-density for the weakest meteors observed as $1{\cdot}07 \times 10^{12}$ electrons/m.

An expression relating electron density to meteor magnitude has been given by KAISER (1955) and is

$$M = 40 - 2{\cdot}5 \log q$$

in m.k.s. units, q being the electron line-density. This gives the minimum meteor magnitude observed on the Gibraltar–Winkfield path at 37 Mc/s as 10. So the oblique incidence results are concerned with smaller meteors than are contained in the Jodrell Bank distribution. It is, of course, possible for the smaller meteors to have a different radiant distribution from the larger ones, and this would account for the failure of the Jodrell Bank distribution to give a good fit to the azimuth results.

It is interesting to note that a similar conclusion about the satisfactory fit of the uniform heliocentric distribution has been reached as a result of Norwegian work (ENDRESEN *et al.*, 1958) on this subject.

The observations on 70 Mc/s have also been compared with the calculated

VOL. 17 1959/6

azimuth distributions based on the same transmitter polar diagram as for 37 Mc/s. In this case, however, the fit is nowhere good. The total values of w^2 obtained were as follows.

Table 5. Goodness of fit of azimuth distributions *A*, *B* and *C* to the 70 Mc/s results

Distribution	Total value of w^2 (hr^{-2})
A	5460
B	4500
C	5410

Because of the poor fit and the absence of experimental data about the transmitter aerial polar diagram, it is not proposed to discuss these results further. It may be mentioned, however, that the fit with distribution *C* can be considerably improved by using a transmitter polar diagram centred on the great-circle bearing and by making it narrower.

5. Conclusions

The azimuth distribution of meteor reflections obtained on 37 Mc/s from Gibraltar is best fitted by a distribution derived from a uniform heliocentric distribution of sporadic meteor radiants. It appears that this should form a satisfactory basis for calculations on forward-scattering by meteors on radio links. The other radiant distributions considered in this paper, which do not give such a good fit, are a uniform heliocentric distribution confined to the plane of the ecliptic, and a distribution of sporadic meteors of magnitude about 7 prepared by the workers at Jodrell Bank. It is also concluded that meteors as weak as magnitude 10 could be observed on the Gibraltar–Winkfield path.

Azimuth results on the 70 Mc/s transmissions from Gibraltar were similar to, but not identical with, those obtained by Meadows on 37 Mc/s. A difference in the transmitter aerial polar diagram in the two cases might account for this.

Acknowledgement—This work was carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research, and this note is published by permission of the Director of Radio Research.

References

Bain W. C. 1958 *Proc. Instn. Elect. Engrs.* B **105**, Suppl. No. 8, 53.

Crow D. A., Kitchen F. A., Isted G. A. and Millington G. 1956 *Nature, Lond.* **178**, 1280.

Davies J. G. 1957 *Advances in Electronics and Electron Physics.* Vol. 9, p. 95. Academic Press, New York.

Endresen K., Hagfors T., Landmark B. and Rödsrud J. 1958 *J. Atmosph. Terr. Phys.* **12**, 329.

FORSYTH P. A., VOGAN E. L., HANSEN D. R. and HINES C. O.	1957	*Proc. Inst. Radio Engrs., N.Y.* **45**, 1642.
HINES C. O.	1955	*Canad. J. Phys.* **33**, 493.
HINES C. O. and PUGH R. E.	1956	*Canad. J. Phys.* **34**, 1005.
KAISER T. R.	1955	*Meteors.* Special Suppl. to *J. Atmosph. Terr. Phys.* **2**, 55.
KENDALL M. G.	1948	*The Advanced Theory of Statistics.* Vol. II. Griffin, London.
KENT G. S.	1958	*Proc. Instn. Elect. Engrs.* B **105**, Suppl. No. 8, 65.
LOVELL A. C. B.	1954	*Meteor Astronomy.* Oxford University Press.
MEADOWS R. W.	1958	*Proc. Instn. Elect. Engrs.* B **105**, Suppl. No. 8, 56.
WEISS A. A.	1957	*Aust. J. Phys.* **10**, 77.
WILLIAMS D.	1958	*Proc. Instn. Elect. Engrs.* B **105**, Suppl. No. 8, 19.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 205 to 219. Pergamon Press Ltd. Printed in Great Britain

Resonance scattering by atmospheric sodium—VII Measurement of the vertical distribution in twilight*

H. N. RUNDLE, D. M. HUNTEN and J. W. CHAMBERLAIN
University of Saskatchewan, Saskatoon and Yerkes Observatory, University of Chicago

(*Received* 14 *May* 1959)

Abstract—The difficulty of determining the distribution of sodium from the usual integral equation for the brightness of twilight is discussed. It is found that the first derivative of this equation is much more suitable. An approximate solution in matrix form is derived and tested. The effect of resonance absorption on such calculations is considered and found to be slight for the lower range of total abundance encountered in practice, but not for the higher. A method of measuring the derivative with respect to shadow height of the twilight brightness is described: the field of view of a birefringent photometer is rotated around the zenith in a small circle and the resulting modulation recorded. Electrical compensation makes the instrument highly insensitive to scattered white light. The twilight emission may be followed right into the night airglow. Distributions are found for four runs made under rather poor conditions. They are similar to previous results but show more detail.

1. INTRODUCTION

OBSERVATIONS of the resonance scattering of sunlight by sodium atoms in twilight offer the possibility of determining the vertical distribution of the sodium. If this could be done with reasonable accuracy, the results would throw light on some of the puzzling variations in sodium abundance, both seasonal and random. If the shadow of the earth were sharp, the distribution could be determined with little difficulty, but this is far from being true; both the lower atmosphere and the ozone layer make important contributions to the shadow and in fact contribute more to the shape of the twilight intensity curve than does the distribution of the sodium. But even so, it has been possible to show that the sodium is mostly in a fairly thin layer near 85 km. Among the first measurements of this type were those made by BARBIER (1948); he was able to fit them with several model layers, but the most plausible had a sharp cutoff at the bottom of height 70 km, and above this an exponential decrease of scale height 8 km. Photoelectric detection allowed HUNTEN and SHEPHERD (1954) to improve the precision somewhat and conclude that a more gradual cutoff at the bottom gave a better account of the observations; their model layer was symmetrical with a scale height of 7·5 km above and below the peak, which was found to be at 85 km. The model used by BLAMONT *et al.* (1958) was Gaussian, with the half-thickness found from the observations; its average value was 14 km and the height of the peak 88 km; these agree well with the earlier values.

A study of the seasonal variation of height was made by HUNTEN (1956a); the symmetrical exponential layer was assumed and only height and total abundance measured. The variation in average height was found to be 1 km or less, after

* The research reported in this paper was supported by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command. The instrumentation, observations, and analysis (Sections 2, 4, 5 and 6) were carried out by RUNDLE and HUNTEN under Contract AF 19 (604)–1831 and with the aid of two Studentships from the National Research Council. The resonance–absorption correction (Section 3) was developed by CHAMBERLAIN and HUNTEN and supported also by Contract AF 19 (604)–3044.

allowance had been made for variations in ozone. It is surprising that the large seasonal variation in abundance is not accompanied by any appreciable change in mean height; this observation puts a severe restriction on any theory of the variation. More detailed observations of the vertical distribution would contribute to the solution of this puzzle.

All these previous investigations were made by measuring the intensity of the *D*-lines as a function of solar depression, and all agree in finding that these measurements carry little information about the sodium distribution. There are two main reasons for this: as already mentioned, the shape of the shadow is important; and the observations must be in effect differentiated to find the distribution. The purpose of this paper is to report a method of measuring the derivative without the use of the integral curve, a method of analysing such observations, and a few preliminary results. The first difficulty still remains, and it is likely that the uncertainty in the transmission function (or shadow shape) now gives most of the error in the results. The apparatus is a birefringent photometer whose output is proportional to the brightness of the sodium emission; its field of view is rotated about the zenith in a circle whose radius is about 5°. The signal is thus modulated by an amount proportional to the difference in *D*-line brightness towards and away from the sun; this is the required derivative modified by two factors which are easily calculated. The modulation is detected and recorded along with the total signal which is proportional to the total brightness. The total brightness includes a "double-scattering" component which can be disturbing towards the end of twilight; light scattered by sodium at the horizon is scattered into the photometer by the lower atmosphere just above it. The derivative arrangement is insensitive to this component since its intensity will vary across the zenith by very little. The idea of oscillating the field of view was taken from a paper by Karandikar (1955) who used it with total light to search for discontinuities in the atmosphere.

To find the distribution of sodium from the derivative, it is necessary to solve an integral equation. This was done by replacing it with a set of linear equations and inverting the matrix of the coefficients. It was found that this would not work if the number of equations was greater than five; thus five points on the distribution were found. Another set may be found by displacing the matrix slightly; the ten points define the distribution fairly well. The results found so far are in reasonable agreement with earlier ones, but show more detail, of course.

When the sodium abundance is fairly great, as in winter time, the brightness is diminished by resonance scattering of the incident sunlight on its first passage through the sodium layer and is further affected by multiple scattering in the neighbourhood of the observed region. The magnitude of this radiative-transfer effect varies with the angle of solar depression and requires an appreciable correction to the observations before they can be used to find a vertical distribution. Previous papers in this series (concerned with early twilight, the nightglow, and the dayglow) have treated this problem with the simplifying assumption of a plane-parallel atmosphere that is illuminated uniformly. But the height distribution necessarily requires observations made when the earth's shadow is crossing the line of sight in the sodium layer. Consequently no simple method exists for computing the multiple scattering into the line of sight.

VOL.
17
1959/6

In this paper we have therefore considered only resonance absorption—that is, we have neglected secondary and higher order scatterings, whose contribution is small in any event. And with the derivative photometer the error thus introduced is even less, since the contribution from secondary scatterings would be expected to have less relative variation over the field of scan than does the primary component. A more accurate allowance for secondary scattering could be accomplished by techniques used by DONAHUE and his collaborators, wherein the radiation falling on a given element of volume is computed by considering the primary scattering in all the surrounding elements of volume, as well as the direct attenuated sunlight. This method is so complex as to require, however, very lengthy calculations with high-speed computers for any assumed model of the sodium distribution. Hence we have proceeded with our admittedly approximate correction.

2. MATRIX SOLUTION OF THE INTEGRAL EQUATION

As far as possible the notation will follow that used by HUNTEN (1954). The transmission of the lower atmosphere for the incident sunlight is represented by $T(z-z_1)$, where z is height and z_1 is the height of the geometrical shadow in the direction observed. Because of atmospheric refraction, some light appears below z_1 to a distance of about 25 km. It is assumed that T does not change appreciably during the part of twilight under consideration; this is a good approximation except for the effect of refraction, and fortunately the refracted light is so weak that a change in it can contribute very little. Resonance absorption makes the scattered intensity less than it would otherwise be; it is taken into account by a factor $\mathscr{S}(z_1)$ which is a number usually between 0·5 and 1 for the abundances normally found. Its calculation is discussed in the next section. If $n(z)$ represents the number density of sodium atoms, the scattered intensity is proportional to

$$\mathscr{S}(z_1).\int_{z_1-25}^{\infty} T(z-z_1).n(z).dz$$

Let $B_s(z_1)$ represent this intensity divided by $\mathscr{S}(z_1)$ and by all the factors of proportionality; then

$$B_s(z_1) = \int_{z_1-25}^{\infty} T(z-z_1)\,n(z)\,dz \qquad (1)$$

Since we are interested in finding the distribution $n(z)$ rather than the total abundance, the factors do not concern us; their calculation has been thoroughly discussed already.

According to previous results discussed in the Introduction, the forms of B_s and T are such that n cannot be determined with the desired accuracy; this is because solving (1) includes a process which is more or less equivalent to differentiating B_s and this introduces additional difficulties. It is preferable to carry out the differentiation separately by a numerical process, or to measure the derivative directly as our new apparatus does. Let us differentiate (1) with respect to z_1; this is easy because the integrand vanishes at both limits, and the result is

$$\frac{d}{dz_1}(B_s) = \int_{z_1-25}^{\infty} \frac{d}{dz_1} T(z-z_1).n(z)\,dz$$

which will be written for convenience as

$$\dot{B}_s(z_1) = \int_{z_1-25}^{\infty} \dot{T}(z-z_1)\, n(z)\, dz \qquad (2)$$

Equation (2) is a linear Volterra equation of the first kind, and is essentially non-singular since the upper limit may be replaced by a large finite height without really changing the equation. A formal solution is known but it cannot be used because it requires the use of high-order derivatives of both $\dot{B}_s$ and $\dot{T}$. Thus we are forced to use an approximate method of solution, replacing the integral by a summation and the continuous variables by averages over finite height intervals. The result is the system of linear equations

$$\dot{B}_s(i) = \sum_{j=1}^{m} \dot{T}(j-i)\, n(j) \quad (i = 1, 2, \ldots, m)$$

The height interval which should have replaced dz has been absorbed into the numerical factor which is being carried in $\dot{B}_s$. This system may be written as a single matrix equation

$$\dot{\mathbf{B}}_s = \dot{\mathbf{T}}.\mathbf{n}$$

in which $\dot{\mathbf{T}}$ is a square matrix of order m, and the other two are column vectors of the same order. The solution is

$$\mathbf{n} = \dot{\mathbf{T}}^{-1}.\dot{\mathbf{B}}_s \qquad (3)$$

in which $\dot{\mathbf{T}}^{-1}$ is the inverse of $\dot{\mathbf{T}}$. Inverting matrices is easy when an electronic calculator is available; once the inverse has been found, the solution may be produced quickly either by the same machine or by a desk calculator. Our calculations were done by an LGP-30 computer using a standard routine supplied with it.

Even though the solution (3) avoids many of the difficulties of the analytical methods, it must still be used with care. A high order m gives a detailed distribution but requires very high accuracy in the observed $\dot{\mathbf{B}}_s$ and the calculated $\dot{\mathbf{T}}$; if these are not attained the results are meaningless. Instead of a smooth distribution one finds large oscillations, usually large enough to carry alternate values negative. When this happens, the set of equations and the matrix $\dot{\mathbf{T}}$ are said to be "ill-conditioned". This trouble is a reflection of the fact that $\mathbf{n}$ is difficult to determine, especially below the peak, even though (2) is much better than (1). Ill-condition may be avoided by choosing the order very small, say 3, but this does not give enough detail in the distribution. The best choice of order must be accompanied by a suitable choice of height intervals so that the matrix equation represents (2) faithfully enough. After some half-dozen unsuccessful attempts, a solution of the fifth order was found which gave good results on a set of brightnesses synthesized from (1), reproducing the original distribution with errors of a few per cent near the peak. Four of the intervals were 7·5 km high; the fifth (at the bottom) was 15 km. The range of z was 62·5 to 107·5 km, which should include nearly all the sodium. The observations $\dot{B}$ were divided into intervals all 7·5 km high in z_1 (shadow height) from 41·25 to 78·75 km; the strange numbers arise from taking the centres of the intervals at heights 45, 62·5, ... km. The same $\dot{\mathbf{T}}$ may be used if the intervals in both $\dot{\mathbf{B}}$ and $\mathbf{n}$ are raised by half their width, 3·75 km; this second set of results is not entirely independent of

VOL.
17
1959/6

the first if the observations are averaged over the whole of each interval, but it helps to define the distribution curve.

The transmission function $T(z-z_1)$ and the matrices $\dot{\mathbf{T}}$ and $\dot{\mathbf{T}}^{-1}$ depend on attenuation by atmospheric ozone and therefore have a seasonal as well as a random variation. The latter cannot be taken into account, but the former can, and the results give guidance as to how serious is the effect of ozone variations in general. A new calculation of T was made by essentially the method described by HUNTEN (1954) and using the ozone distributions adopted (1956b) for the minimum in the fall and the maximum in spring. The absorption coefficient was increased to 0·059 cm^{-1} (base 10) to agree with the value now generally accepted. Attenuation by haze in the lowest atmosphere was included, using the scale height of 1 km suggested by the observations of NEWKIRK (1956) and a coefficient 0·060 per atmosphere (base e; ALLEN, 1955). The effect is to cut off some of the tail of the transmission below the geometrical shadow; the transmission is 0·1 per cent about 14 km below it. Fortunately, this change has very little effect on the results for the height and distribution of sodium; they depend mostly on the transmission from 15 to 40 km above the geometrical shadow. The final result was two transmission functions for "fall" and "spring". The spring one was used to synthesize the test brightnesses already mentioned, with a double-exponential model distribution. These calculations were done on the electronic computer with an approximation to (1) using about forty ordinates in each integrand. The derivative of the brightness was found by differencing these results. The fact that these test results were calculated with (1) rather than (2) gives a further check to the method of solution.

The transmission derivative $\dot{T}(z-z_1)$ was also found by differencing accompanied by a little graphical smoothing. Though the matrix contains twenty-five components, only thirteen of them are different, and the labour of finding them is not great. The method was tested by applying the spring inverse to the test results, with the satisfactory answer already discussed. To see how great is the effect of ozone variations, the fall inverse was then applied to the same test results; the answer was noticeably worse, but still recognizable. It was concluded that one more matrix intermediate between the original two would cover the year adequately; it would be used for both winter and summer, but was given the designation "winter" for convenience. It was found by averaging the $\dot{\mathbf{T}}$ matrices for spring and fall, and then inverting the average. The winter inverse applied to the test results gave an answer within a few per cent of that found with the spring inverse. The random fluctuations of ozone are considerably less than the seasonal variation which causes the difference between the matrices used; thus it may be concluded that their effect on the solution is negligible.

If these solutions are to be used at other stations, it is recommended that they be considered as representing certain ozone abundances rather than certain seasons. These are 0·213 cm for the "spring", 0·249 cm for the "winter", and 0·285 cm for the "fall". The ozone abundance should be within about 0·02 cm of the value given for reasonable accuracy.

The work described above was completed before the calculations of T by LINK (1958) became available. These remove the approximation that T depends on $(z-z_1)$ alone, but this should make very little difference. On the other hand, we

have included variations in ozone and tried to take into account the absorption by dust and haze; we thus consider our results more useful for the present purpose.

3. The Resonance Absorption Function $\mathscr{S}(z_1)$

If resonance absorption remained constant during a given twilight, it would have no effect on the problem of determining the vertical distribution. In fact, however, it does depend slightly on the solar depression because the slant path through the sodium changes. If this were the only effect, $\mathscr{S}(z_1)$ would increase with z_1 through the whole of twilight as it certainly does until the shadow begins to shade the sodium layer. But from this point on the behaviour is doubtful because another effect enters: the presence of the shadow forces the light to traverse a longer path through the sodium and this tends to compensate for the otherwise reduced slant path. The result is that $\mathscr{S}$ is almost constant through the part of twilight which is useful for determining the distribution of sodium, and the correction to the distribution is small. However, it may still be appreciable, and it is necessary to estimate the variation $\mathscr{S}(z_1)$ as precisely as possible. For our present problem of finding the distribution, absolute accuracy is not as important and in any case good results for early twilight are already available from earlier papers in this series (Chamberlain *et al.*, 1958, Paper IV) and the work of Blamont *et al.* (1958). Donahue (1958) has commented on the problem of calculating $\mathscr{S}$ in late twilight and referred to results of Stull; but we feel that ours are better for the present problem in which smoothness is more important than accuracy.

To a first approximation, resonance absorption in early twilight may be treated by taking an "average" ray which is scattered at the centre of the sodium layer (solid line in Fig. 1). When a shadow is present this path may no longer be available and one like the dotted line in Fig. 1 may be more important. Since the shadow is

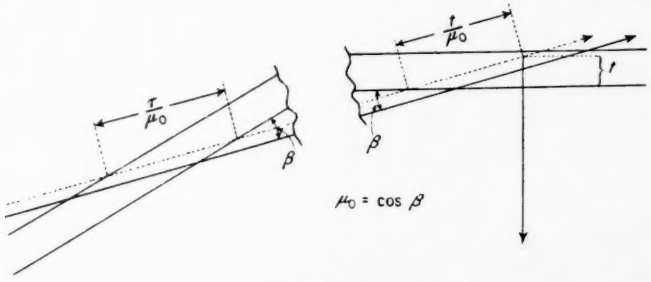


Fig. 1. The passage of sunlight through the sodium layer on the day and twilight sides of the earth. When shadowing is not important, the average ray follows the solid line; when it is, paths such as the dotted line may be preferred.

not sharp it is necessary to average over all possible paths, with each one weighted by the proper $T(z-z_1)$. Further complication enters from the presence of two D-lines, each with hyperfine structure. Therefore the averaging over the Doppler profile of each component is avoided by the use of an effective absorption coefficient equal to the peak value divided by $\sqrt{2}$ (Hunten, 1956c; Brandt and Chamberlain, 1958); and single scattering only is included. These approximations should be acceptable up to abundances of about 10^{10} atoms/cm^2, especially since we require

relative rather than absolute accuracy, and since the photometer tends to ignore doubly-scattered light. In previous papers of this series it was not necessary to assume a distribution of sodium with height, since it has a negligible effect, but here the presence of the factor T requires the use of a distribution reasonably like the actual one. A simple exponential with cutoff at 80 km was used; this should give similar results to a symmetrical layer with maximum at 85 km. This distribution is represented by

$$n(z) = (N/H)\exp[-(z-z_c)/H]$$

where N is the abundance (atoms/cm^2), H is the scale height of the distribution (7·5 km), and z_c is the height of the cutoff (80 km).

For a single Doppler component (i) the vertical optical depth at the centre is $\tau_{0i} = N\alpha_i$ where α_i is the peak cross-section given in Paper IV, Table 1*. The effective optical depth for the whole line is then $\tau_i = \tau_{0i}/\sqrt{2}$. The optical depth up to the height z is called $t_i(z)$ and for the model chosen is

$$t_i(z) = \tau_i\{1-\exp[-(z-z_c)/H]\}$$

Over the dotted path in Fig. 1 the total attenuation is $\exp[-(\tau_i/\mu_0 + t_i/\mu_0 + t_i)]$; this must be weighted by the intensity of scattering at height z which is $T(z-z_1)\,n(z)$. Thus for the single component we have

$$\mathscr{S}_i = \exp(-\tau_i/\mu_0).\frac{\int_{z_c}^{\infty} T(z-z_1)\,n(z)\exp[-t_i(1+1/\mu_0)].dz}{\int_{z_c}^{\infty} T(z-z_1)\,n(z)\,dz}$$

The values for the four components are then averaged with weights proportional to $r_i\alpha_i$ to give the final expression for $\mathscr{S}$; the four terms may be combined since they have a common denominator. r_i is the residual intensity of the appropriate Fraunhofer line. The integrals are similar to the one in (1) and the expression was calculated by a modification of the same computer programme. Differences at 2 km intervals were calculated and printed also; these yield $\dot{\mathscr{S}}$ (i.e. $d\mathscr{S}/dz_1$) which is the real object of the calculation; better values of $\mathscr{S}$ are already available but the older results do not give a useful $\dot{\mathscr{S}}$.

The behaviour of $\dot{\mathscr{S}}$ is somewhat surprising; as expected, it starts to decrease as the shadow reaches the sodium layer, going slightly negative at 60 km; but then it rises to a peak at 77 km and falls to near zero at 87 km, the limit of the calculations. The shape of the curve of $\dot{\mathscr{S}}$ vs. z_1 is very similar for any value of abundance between 1 and 10×10^9 atoms/cm^2.

It is now necessary to consider how the quantity $B_s(z_1)$ appearing in (2) can be found from observations of $4\pi B$ and $4\pi\dot{B}$; however, we continue to omit constants of proportionality which are not convenient to carry when we are interested in the distribution. With this understanding, we have

$$B_s = B/\mathscr{S}, \quad \text{and} \quad \dot{B}_s = \dot{B}/\mathscr{S} - \dot{\mathscr{S}}B/\mathscr{S}^2$$

* Properly speaking one should use $(\varpi_i/\varpi_s)\alpha$ instead of α, but we retain the simpler notation especially since it has already been used in Paper IV. (The ϖ's are statistical weights.)

Now to a good approximation we may consider $\mathscr{S}$ as constant, since its slight variation is taken care of by the second term; the factor $1/\mathscr{S}$ may thus be omitted, and we have

$$\dot{B}_s \doteq \dot{B} - \dot{\mathscr{S}} B/\mathscr{S}$$

Since the second term is a small correction, it does not need to be known with high accuracy, and it should be sufficient to use standard values calculated for an average B; of course, these are still a function of the abundance.

To see the effect of this correction and to get an idea of its magnitude, it was applied to a set of measurements made at an abundance of about 5×10^9 atoms/cm^2. The effect is almost confined to the region below 85 km, and is to move some sodium up from the lowest interval into the second-lowest; the cutoff is sharpened but the centre of gravity moves up only 0·5 km. But the amount of sodium involved is small, about 5 per cent of the total; thus the correction is of the same order as the errors in the measured distribution. It may be concluded that the correction is appreciable only for abundance greater than 5×10^9 atoms/cm^2 unless further improvements are made in the accuracy with which the distribution is measured in the first place. Since the measurements reported here are of lower abundance, no correction for resonance absorption is made to them, except in finding the total abundance.

4. Apparatus

Measurements of sodium in twilight are often made near the horizon at zenith angles near 75° because the intensity there is about three times greater than in the zenith. But this location has serious disadvantages as well; interpretation is made still more difficult, and additional uncertainties in the transparency of the atmosphere arise (Hunten, 1956b; Lytle and Hunten, 1959). Thus it was considered essential to confine measurements to the neighbourhood of the zenith even though greater sensitivity is required of the photometer.

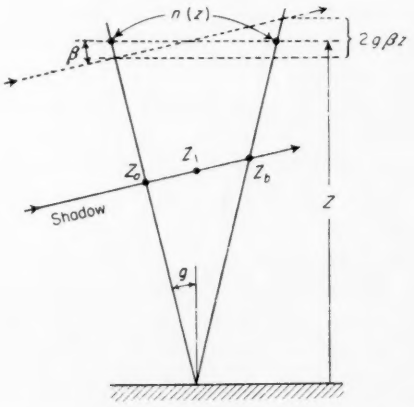


Fig. 2. Schematic illustration of the two effective fields of view of the photometer. The half-angle of oscillation g and the solar depression angle β are both small; the height-difference $2g\beta z$ is about 2 km.

The way in which the apparatus measures the derivative of the brightness may be understood from Fig. 2, which is a section of the earth in the vertical plane containing the sun. The photometer must give a continuous electrical signal proportional to the brightness of the *D*-lines, ignoring the Rayleigh-scattered sunlight, and must have a small field of view, about 1°. The field is oscillated so that it observes alternately two points in the sky near the zenith at equal zenith angles of a few degrees. The earth's shadow is slightly lower on the side nearer the sun, and thus the twilight is in a slightly earlier phase. A switch moving in synchronism with the oscillation of the fields can then be arranged to take the difference between the two signals, and this is almost proportional to the desired derivative. Just what it is will be considered a little later.

For convenience, the field is not actually oscillated as this description assumed; instead, it is rotated in a circle about the zenith, but this motion is essentially equivalent. The motion is produced by a pair of thin prisms of plate glass ("objective prisms") rotating at 30 rev/min and having a total deviation of 4° 40′; the circle described by the field of view has a diameter of 9° 20′. Two prisms were used instead of one so that smaller angles of deviation would be available; but experience showed that the largest angle was necessary, and a still larger one might be useful.

Just below the prisms is a plano-convex lens of the same diameter, 30 cm, and a focal length of 230 cm; at the focus a 6·4 cm diaphragm defines the field of view, about 1·6°. The components are mounted in a square wooden tube which is fixed in a vertical direction but able to rotate about its axis so that the sun's azimuth can easily be followed around the horizon. This part of the system is a vertical twilight telescope for monochromatic light, able to collect a large amount of light from a small field of view.

The photometer itself is mounted at the bottom of the telescope; as already mentioned, it must be able to produce a signal which is independent of the white component of the twilight. A second requirement is high sensitivity so that the twilight can be followed right down to the intensity of the night airglow, and so that the airglow intensity can be measured. Unfortunately, the Zeeman-modulated sodium cell of Blamont (1956) is not sensitive enough, although it is almost ideal for the bright part of twilight. A similar cell used in transmission instead of scattering would probably serve, but this has never been tried. The birefringent filter of Blamont and Kastler (1951) has high sensitivity and was the one used; it has more recently been applied to studies of the night airglow by Dunn and Manring (1946) and others. Blamont and Kastler found that the photometer was disturbed by the Fraunhofer lines in the white light, but we have been able to reduce this effect greatly by a simple compensation procedure.

The birefringent element was made of ADP (ammonium dihydrogen phosphate) instead of quartz because it was quickly available; but experience has shown it to be greatly inferior because of its large temperature coefficient. The temperature must be regulated to 0·01°C and this is difficult; in practice adjustments had to be made every half-hour or so. The Polaroid analyser rotating at 27·5 rev/sec modulates the emission and absorption lines at 55 c/s; in practice only the *D*-lines enter the system because an interference filter with a bandwidth of 10 Å is included. When this

signal has been detected by a phase-sensitive detector, the emission lines give a positive signal and the undesired Fraunhofer lines a negative one; the latter must be eliminated. This is done by making use of the unmodulated component of the photomultiplier signal, which depends mostly on the white light; a suitable fraction of this is added to the detected signal until the response to attenuated daylight is zero. The procedure is discussed in more detail by Donahue and Hunten (1958). The procedure works well and allows measurements to be made for shadow heights z_1 greater than 35 km or solar depressions greater than 5°. This more than covers the region of interest for measuring the vertical distribution of sodium.

The detected compensated signal still includes a small modulation, about 10 per cent, at a frequency of 30 c/min, due to the rotation of the prisms. This is detected by a second phase-sensitive detector using an ordinary relay. This circuit gave considerable trouble because earlier versions were sensitive to changes of the signal in time as well as the desired changes in space. A well-balanced circuit was finally devised which had the required properties; it was tested and gave no signal in a run with the two prisms opposing so that the field of view was fixed in the zenith. Both the total signal and the difference one are recorded on a two-pen chart recorder.

Because of its strong discrimination against white light, the instrument must be calibrated with a source incorporating a sodium lamp. The standard which was constructed contains a vacuum phototube to measure a small fraction of the light abstracted from the beam by an unsilvered piece of glass. A large Fresnel lens and a sheet of opal glass give a nearly uniform illumination over a 12 in. circle. A series of known apertures can be placed over another diffuser near the lamp so that known intensity ratios can be produced. A solenoid-operated shutter carrying a calibrated screen can be oscillated into and out of the beam in synchronism with the prisms so as to simulate the small intensity differences observed in twilight. With this calibrator the photometer was tested and found linear in both channels within the accuracy of measurement.

The absolute brightness of the calibrator was measured by comparing it with the standard low-brightness source (Shepherd, 1954) by means of the scanning spectrometer previously used for twilight work. The measurement allowed for the fact that the edges are not quite as bright as the centre. Direct comparison on a fairly bright twilight measured with both instruments would give a useful check; a few tests made on faint twilights showed no disagreements greater than the 10 per cent uncertainty from the low sensitivity of the spectrometer. Comparisons of this sort show dramatically how much better the photometer is than the spectrometer; the latter can barely measure the sodium in summer, but the photometer gives an excellent signal and can follow it right into the night airglow.

Because of the large size of the beam of light in this instrument, the photomultiplier used for the measurements reported here was one with a 2 in. diameter cathode, and the dark current was high even at −78°C; this meant a correspondingly high noise level. More recently an EMI 9502B has been installed; the cathode is only 1 cm in diameter and it was necessary to cement an aplanatic lens to the front to condense the beam enough. The dark current at −78° is only about 1 electron/sec, and a mechanical refrigerator producing a temperature of −45° has replaced the

VOL.
17
1959/6

dry ice; the dark current is thus about 5 electrons/sec and the signal-to-noise ratio highly satisfactory.

It has already been mentioned that the quantity measured by this apparatus is not quite the derivative $\dot{B}$; instead it is a related quantity which we will call ΔB. Let g represent the effective angle of oscillation as in Fig. 2; this will be somewhat less than the half-angle of the conical scan. Consider the light scattered by the two unit volumes shown at height z; the contribution to ΔB from these is

$$\begin{aligned} d(\Delta B) &= [T(z-z_a) - T(z-z_b)]\, n(z).dz \\ &\doteqdot \dot{T}(z-z_1).2g\beta z.n(z).dz \end{aligned}$$

$$\Delta B \doteqdot 2g\beta \int_{z_1-25}^{\infty} \dot{T}(z-z_1).zn(z).dz \qquad (4)$$

This equation is similar to (2) and may be solved in the same way; if ΔB is divided by the solar depression angle β and corrected if necessary for resonance absorption, it may be used in (3) with the same matrix and will yield $zn(z)$ instead of $n(z)$. It is possible to work out another matrix incorporating the variations of β and z and thus get $n(z)$ directly with only the correction for resonance absorption; this may turn out to be best for routine calculation, but for the present we prefer to retain the other procedure since it is more flexible; in particular it is better suited to finding two sets of solutions shifted by half an interval. It may also be used with a derivative found from the total intensity curve.

In the following we call ΔB the "derivative" with the understanding that the term is used for convenience and must not be taken literally.

The method we have described for measuring the "derivative" directly in twilight has several advantages, but of course some disadvantages as well. The advantages are: under favourable circumstances the direct measurement is more accurate than differentiation of the total intensity curve; multiple scattering in both the sodium layer and the lower atmosphere has little effect; and slow variations in atmospheric transparency are not important. The most important disadvantage is that serious errors may arise because the sodium layer is not horizontally uniform. These may be reduced by keeping the angle of scan small, but at the expense of signal strength; it appears that the angle used is as small as practicable. The effect of non-uniformities can be assessed only by experience; our results so far seem to indicate that it is usually small. Variations at right angles to the sun's plane could be measured by a second phase-sensitive switch arranged to take the difference between the two sides of the conical scan; this would not give information directly on conditions in the orthogonal direction for a single twilight, but would give useful statistical information from which probable errors could be deduced. We have not yet included such a device because it requires an extra recorder and some other apparatus, but it is well worth considering.

A much more detailed description of the apparatus and the mathematics is given by Rundle (1958).

5. Observations

The four sets of results reported are the best obtained during the first 2 months' operation of the new instrument, and should not be taken as demonstrating its best

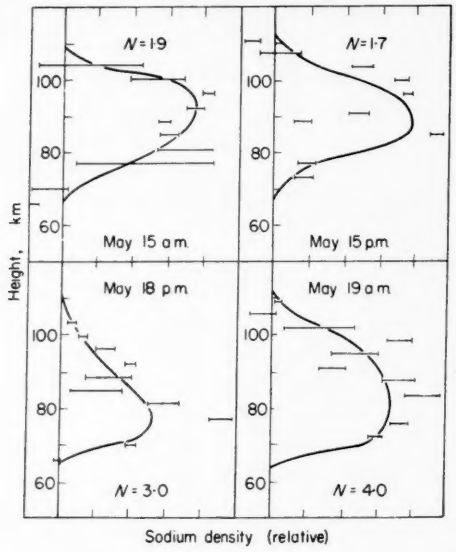


Fig. 3. The four measured distributions; the errors shown with the points are derived from the estimated errors of observation so that the total error is probably greater. Total abundances N in units of 10^9 atoms/cm^2 are shown.

VOL.
17
1959/6

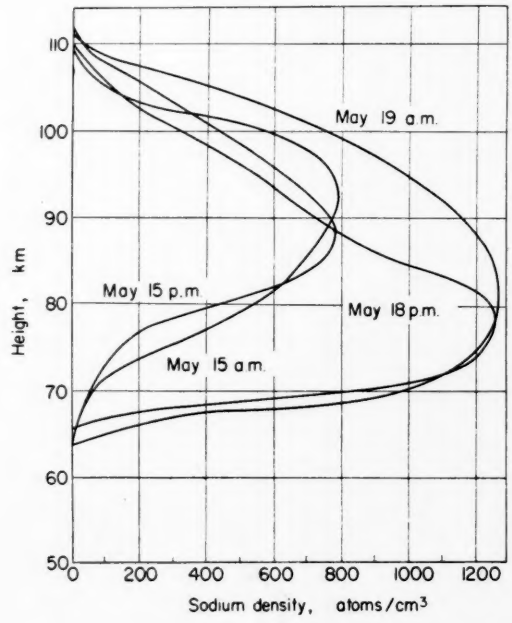


Fig. 4. The four distributions of Fig. 3 plotted together for comparison.

capabilities. There are two reasons for this: the intensity of the sodium twilight was near its seasonal minimum and the four brightest twilights were used; and the photometer had just been put into operation and was using a noisy photomultiplier. The rest of the runs gave no "derivative" signal that could be seen in the noise, even with a filtering time-constant of 30 sec, and these four gave signal-to-noise ratios for ΔB of only 1 or 2. Much better results may be expected when the instrument is operating properly in its final form. But even with these limitations the distributions obtained are an improvement over those found by previous methods.

Because of the low signal-to-noise ratio, the traces were measured with a planimeter to give the average deflexion for each required interval; a smooth curve was then drawn through the resulting points. This curve was compared with one drawn by eye through the noisy trace; the agreement was satisfactory. Two column matrices of "$\dot{B}$" were derived, shifted by half an interval as described in Section 2. No correction for resonance absorption was necessary since the abundances were less than 4×10^9 atoms/cm^2. Multiplication by the inverse matrix T^{-1} gave the two sets of $zn(z)$; the $n(z)$ from these are plotted in Fig. 3. Because of the linearity of the matrix equations, the effect of errors of observation may be found by a similar calculation; an error vector is estimated and multiplied by the inverse matrix. These errors are shown with the points in Fig. 3; there are probably other errors which will be considered in the next section. In Fig. 4 the area under each distribution curve has been adjusted in accordance with the total abundance found from the plateau of the total intensity record.

It will be noticed that there is a strong tendency for the points to be alternately above and below the smooth curve; this oscillation seems to be characteristic of matrix solutions. (With larger matrices it is bad enough to mask the distribution completely.) The fact that the points oscillate in pairs results from the superposition of two matrix solutions. Thus we feel justified in drawing smooth curves between the points. The distributions are reasonable and similar to those deduced by other methods. They were checked by integrating back with the computer programme for equation (1) and the results compared with the total intensity curve observed; the agreement was reasonable. Some of our arguments for the superiority of the "derivative" observations suggest that the agreement should not be perfect; in other words, our synthesized curve does not include the multiple scattering which affects the observed one. It may be that some of the discrepancies are caused by horizontal variations in the sodium layer which may have a large effect on the difference measurements.

6. Discussion and Conclusions

The four distributions are plotted together in Fig. 4. It is evident that the two pairs which were observed in successive twilights are similar, but that there has been a large change in the $3\frac{1}{2}$ days separating the two pairs. The agreements are our best evidence that the measurements are fairly good and give a reasonable estimate of the possible errors. It is striking that the doubling of the total abundance seems to have been caused entirely by sodium at the bottom of the distribution; whether this is typical cannot be decided without much more observation. This is

an example of the possible advances in our knowledge which may come from detailed distribution measurements.

The effect of errors of measurement is shown on Fig. 3, but there are other possible errors which may be more serious. The matrix analysis is not above suspicion, but the tests described in Section 2 seem to show that it is adequate as long as the correct transmission function is used. But there is no way of knowing whether this is correct or whether it varies appreciably from day to day. Certainly it does vary; the ozone abundance and distribution seem to change, and the attenuation of the lowest atmosphere is highly variable. It seems likely that the oscillations seen in Fig. 3 are the result of this difficulty; but even in their presence it is still possible to draw a reasonable curve. Horizontal variations in the sodium layer and the rest of the atmosphere do not seem to have had much effect in these cases. The importance of these factors can best be found by considering a large number of observations.

This paper has described a method of finding the vertical distribution of atmospheric sodium from the derivative of the usual curve of intensity vs. shadow height. The particular method used to solve the integral equation was to approximate it by a set of linear equations and solve them by matrix inversion. The results are fairly good, but it might be profitable to look for other methods of solution. A single application of the matrix gives five points on the distribution; a second application can be used to get five more. There is a tendency for the points to oscillate about a smooth curve, but the latter can still be drawn without much difficulty.

This analysis is general and need not be confined to measurements made with the new photometer. It is possible to differentiate the ordinary integral curve found with a spectrometer or other photometer and still have a derivative of useful accuracy. Even though individual points may show a large scatter, a smooth curve drawn through them can be used to give the data for the matrix solution. One attempt of this sort has been made with encouraging results, and more will be made in the future.

The variation of resonance absorption through twilight must be removed from the measurements before they can give a correct distribution. Previous calculations have not given this variation with adequate precision, although their absolute values have been better. A new calculation was made with the aid of two approximations; single scattering only was included; and an effective absorption coefficient was used for each Doppler line. These should not cause serious errors, especially with the new photometer which tends to ignore multiple-scattered light. When the correction was applied to a set of observations, it was found to have only a small effect for an abundance of 5×10^9 atoms/cm^2, of the order of the errors of measurement. It should thus be negligible for smaller abundances, and was not included in the treatment of the preliminary observations.

A birefringent photometer of the type originated by Blamont and Kastler has been incorporated in a twilight telescope which allows a quantity similar to the above derivative to be measured. The apparatus is still being improved in detail, but enough results have been obtained to show that the principle is sound. Of these first observations, four were bright enough to give a useful difference signal, and the results are presented in Figs. 3 and 4. When the photometer is operating at its best,

VOL.
17
1959/6

it is expected to give a great deal of new information on the vertical distribution of sodium and its variations. Since it can observe the sodium in the night airglow as well, there is hope of being able to study the relationships between the night and twilight glows of the *D*-lines.

Acknowledgements—The first calculations of $\mathscr{S}(z_1)$ were made by J. C. BRANDT; even though they were not finally used, they were useful as a check on the computer programme and in finding faults. Some of the observations were made by B. G. PRATT, who also did most of the work of reducing them to numbers. The patient machine work of F. RITTMANN is also gratefully acknowledged.

REFERENCES

ALLEN C. W.	1955	*Astrophysical Quantities*, p. 116. Athlone Press, London.
BARBIER D.	1948	*Ann. Géophys.* **4**, 193.
BLAMONT J. E.	1956	*The Airglow and the Aurorae* (Edited by ARMSTRONG E. B. and DALGARNO A.), p. 99. Pergamon Press, London.
BLAMONT J. E. and KASTLER A.	1951	*Ann. Géophys.* **7**, 73.
BLAMONT J. E., DONAHUE T. M. and STULL V. R.	1958	*Ann. Géophys.* **14**, 253.
BLAMONT J. E., DONAHUE T. M. and WEBER W.	1958	*Ann. Géophys.* **14**, 282.
BRANDT J. C. and CHAMBERLAIN J.W.	1958	*J. Atmosph. Terr. Phys.* **13**, 90.
CHAMBERLAIN J. W., HUNTEN D. M. and MACK J. E.	1958	*J. Atmosph. Terr. Phys.* **12**, 153.
DONAHUE T. M.	1958	*Phys. Rev. Letters* **1**, 287.
DONAHUE T. M. and HUNTEN D. M.	1958	*J. Atmosph. Terr. Phys.* **13**, 165.
DUNN R. B. and MANRING E. R.	1956	*J. Opt. Soc. Amer.* **46**, 572.
HUNTEN D. M.	1954	*J. Atmosph. Terr. Phys.* **5**, 44.
	1956a	*The Airglow and the Aurorae* (Edited by ARMSTRONG E. B. and DALGARNO A.), p. 114. Pergamon Press, London.
	1956b	*Ibid.* p. 183.
	1956c	*J. Atmosph. Terr. Phys.* **9**, 179.
HUNTEN D. M. and SHEPHERD G. G.	1954	*J. Atmosph. Terr. Phys.* **5**, 57.
KARANDIKAR R. V.	1955	*J. Opt. Soc. Amer.* **45**, 389.
LINK F.	1958	*Studia Geoph. et Geod.* **2**, 47.
LYTLE E. A. and HUNTEN D. M.	1959	*J. Atmosph. Terr. Phys.* **16**, 236.
NEWKIRK G. A.	1956	*J. Opt. Soc. Amer.* **46**, 1028.
RUNDLE H. N.	1958	Ph.D. Thesis, University of Saskatchewan.
SHEPHERD G. G.	1954	Scientific Report AR–16, Contract AF 19 (122)–152, University of Saskatchewan.

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 220 to 245. Pergamon Press Ltd. Printed in Great Britain

Thermal and gravitational atmospheric oscillations—ionospheric dynamo effects included

M. L. WHITE
U.S. Department of Commerce, National Bureau of Standards,
Boulder Laboratories, Boulder, Colorado, U.S.A.

(*Received* 19 *May* 1959)

Abstract—The resonance theory of gravitational and thermal oscillations in a rotating atmosphere composed of a neutral gas (SEN and WHITE, 1955; WHITE, 1956) is extended to include an electron and positive ion gas with a permanent magnetic field superposed (so-called dynamo effect, CHAPMAN and BARTELS, 1940). Basic equations of energy and motion are given, resulting, for example, in a Kirchhoff's voltage law for the upper atmosphere. Height dependent expressions are obtained for the electric field, the current density, the ion drift velocities and non-linear heat source functions $\bar{q}_w$ $\bar{q}_n$ for the gas-as-a-whole and the neutral gas component. The $\bar{q}$-functions are produced by an I^2R power loss. The conditions under which one obtains the previous time-independent differential wave-equation of WILKES (1951) and SEN and WHITE (1955) are discussed; as an example, it is required that the thermal source function $\overline{Q}$ include a $\bar{q}$ function, resulting in two coupled wave-equations, one for the gas-as-a-whole and another for the neutral component.

Using a method of operators, a time-dependent differential wave-equation is obtained from the time-independent wave-equation. It is found as a sufficient condition that if the four independent space and time variables in the non-linear $\bar{q}$-term are separable [with further minor restrictions given in equation (48)], then the oscillations satisfy the condition for neutral stability. Finally, a non-equilibrium "solution" for formation of ionized layers, based on the electron continuity equation, "tidal" transport terms included, is illustrated, using a method of successive approximation. Cavendish tables (THOMAS and ROBINS, 1955) of the daily electron density variations at given true heights and the New York University analysis (KERTZ, 1956a) of the daily surface pressure variations are examined for predicted non-linear effects. One interesting consequence is an explanation of the major contribution to the 6 hr surface pressure variation as a self-coupling of the important 12 hr pressure variation.

VOL.
17
1959/6

1. INTRODUCTION

THE existence of electric currents above the earth was suspected as long ago as 1832, when FARADAY (1832) suggested them as a cause of auroras. STEWART (1882) suggested such currents as the cause of the magnetic daily variation and advanced the dynamo theory which was subsequently elaborated upon by Schuster and Chapman (and summarized by CHAPMAN and BARTELS, 1940), MCNISH (1939), and others. In a sense, the dynamo theory may be considered an extension of tidal theory of a neutral gas by including the addition of small quantities of charges whose motions are described in the presence of the earth's permanent magnetic field. It seems, therefore, appropriate to re-examine the dynamo theory in the light of current advances in the resonance theory of atmospheric oscillations, including its extension to the upper atmosphere as given by the writer (WHITE, 1956).

Some specific advantages of a tidal–dynamo approach would be the possibility of determining how the current systems vary with height, as well as the possibility of eliminating certain assumptions made in earlier discussions of dynamo theory, namely:

(1) Incompressibility of the conducting fluid or plasma through the introduction of a current function.

(2) The neglect of Coriolis forces or the earth's rotation through the introduction of a velocity potential.
(3) Zero vertical velocity.
(4) Plane-parallel ionized layers.
(5) Isotropic conductivity; inclusion of anisotropic conductivity effects have been considered by BAKER and MARTYN (1953) and by CHAPMAN (1956).

In Section 1.2 the ionosphere is considered to be a ternary gas consisting of a neutral gas and an electron and positive ion gas. The basic equations of motion and energy are set up for each of the three components, including effects of a permanent external magnetic field. However, as nearly all of the elementary particles comprising the ionosphere are neutral, F-region included, the thermal and gravitational oscillations are considered as occurring in the neutral component alone. Then through collisions, the neutral particles (in a dynamo action) drive the electrons and positive ions differentially across the magnetic lines of force, thus generating a current. Under these conditions, the equations of motion for the charged particles reduce to a purely electrical law, namely, Kirchhoff's voltage law.

A new term, an $I^2 R$ loss, appears in the conservation of energy (or power) equation, showing the degradation of tidal energy into an electrical heating loss. This loss, arising from collisions of the charged and the neutral particles, constitutes an electrodynamic viscosity, as opposed to collisions of neutrals with neutrals which constitute ordinary dynamic viscosity. In Section 2 it is shown that this non-linear $I^2 R$ term can actually be considered as augmenting the "solar" heating term $\bar{Q}_r$ (forcing function) in the energy conservation equation. The energy conservation equation then takes the same form it had in previous tidal treatment, as do all the other "tidal" equations [after certain limiting assumptions, equation (14)]. By expanding all the time-dependent variables in Fourier series, one obtains an infinite set of linearized "tidal" equations and finally recovers the original differential equation of motion (or wave-equation) of "tidal" theory. In fact, two differential equations of motion obtain, one for the gas-as-a-whole and another for the neutral particles, each differing from the original "tidal" equation by non-linear $I^2 R$ terms, which serve to couple the two differential equations.

In Section 2.1 an operator method is employed to derive a time-dependent wave-equation. This differential equation quickly leads to an expression for the steady winds (i.e. seasonal winds) in terms of a seasonal heating function.

In Section 3 it is shown how the vector drift velocities of the charged particles, the vector conduction current, the electric field vector, and the non-linear $\bar{q}_r$ terms (which involve the vector drift velocities of the charged particles) can be obtained in terms of the wave-functions for the neutral gas and the gas-as-a-whole. This dependence of the q-term in the wave-equation on the wave-functions themselves necessitates a method of successive approximation for solving the differential wave-equations.

Section 4 is concerned with the stability criterion for the non-linear vibrations, produced either by collisions of neutrals with neutrals or neutrals with charged particles. A sufficient condition for neutral stability is that the variables in the non-linear term be separable.

The non-linearity implies a rich harmonic content for the pressure and density variations. A test for the existence of a rich harmonic content is attempted in

Section 5. The predicted non-linearities seem to account for the dominant term of the 6 hr pressure variation at the ground as a self-coupling of the 12 hr vibration, due to ordinary dynamic viscosity. Similarly the large harmonic content of the "electron density vs. time variation" in equatorial regions (at 280–300 km) found from Cavendish tables (THOMAS and ROBBINS, 1955) seems consistent with the theory, the high harmonic content arising in part from the I^2R electrodynamic friction term. Included in Section 5 is recognition of the role of temporal electron density variations in the dynamo theory, through consideration of the electron continuity equation, which includes a "tidal" transport term and a solar ionization term. The solution is indicated not only for the steady state, but for the first and second harmonics of the temporal variations of electron density. The differential equation of electron continuity plus the differential wave-equations for the neutral gas and the gas-as-a-whole form the basic simultaneous non-linear differential equations for the tidal–dynamo problem; their numerical solutions will not, however, be considered here.

It is possible that dynamo effects may become more than a perturbation to the resonance theory: the difference between the tidal velocities of the neutral gas and the gas-as-a-whole could grow appreciably as electro-dynamic viscosity grows in importance. Conceivably these viscosity effects could lead to improved boundary conditions at the "top" of the atmosphere by damping out "tidal" oscillations. Finally, as in previous "tidal" work which involves a "solar" heating function or source function (WHITE, 1956), a model for the latter is necessary before numerical results can be given. Measurements of the daily temperature changes in the upper atmosphere above even one given station can give the needed model.

VOL.
17
1959/6

1.1 *Notation*

The notations used for atmospheric tides (WILKES, 1951; WHITE, 1956) and for atmospheric electromagnetic phenomena (CHAPMAN and COWLING, 1952) overlap. To minimize confusion, the symbols used in the present paper are listed as follows:

a = radius of earth;

ω = angular velocity of the earth;

g = acceleration due to gravity;

C_v, C_p = specific heats of the gas composing the atmosphere at constant volume and at constant pressure;

$\gamma = C_p/C_v$;

R_r = gas constant for unit mass of gas r, where henceforth, $r = e, i, n, w$, i.e. electrons, positive ions, neutral atoms or molecules, and the gas-as-a-whole;

$R \simeq R_w \simeq R_n$;

θ, ϕ, z = co-latitude, longitude, height (above surface of earth) of a point in the atmosphere;

u_r, v_r, w_r = southward, eastward, upward components of "tidal" air velocity at (θ, ϕ, z) for the rth gas component;

$\bar{\chi}_r(\theta, \phi, z, t)$ = divergence of velocity for gas component r; t denotes the time; if only a single harmonic is involved, then a further subscript ν is added, thus:

$\bar{\chi}_{r\,\nu}(\theta, \phi, z, t) = \chi_{r,\nu}(\theta, \phi, z)\,\chi_\nu(t) = \chi_{r,\nu}(z)\,\chi_{r,\nu}(\theta, \phi)\,\chi_\nu(t)$ if the variables are separable;

$\bar{p}_r, \bar{\alpha}_r, \bar{T}_r, \bar{\rho}_r, \bar{n}_r$ = total pressure, mass density, absolute temperature, charge density, number density for gas component r; it is assumed that all the physical variables can be expanded in terms of a Fourier series, for example:

$$\bar{p}_r = \sum_{\nu=0}^{\infty} \bar{p}_{r,\nu} = \sum_{\nu=0}^{\infty} [\bar{p}_{r,\nu}(\theta,\phi,z)\, e^{i\sigma_\nu t} + p_{r,\nu}{}^*(\theta,\phi,z)^{-i\sigma_\nu t}] \tag{1}$$

where $p_{r,\nu}{}^*$ denotes the complex conjugate of $p_{r,\nu}(\theta,\phi,z)$; for $r = n, w$, it is assumed that the static terms are very much greater than the time-dependent (or perturbed) terms, and that they are functions of z only;

σ_ν = the angular frequency of oscillation for the νth harmonic ($\sigma_0 = 0$);

$p_r, \alpha_r, T_r, \rho_r, n_r$ = departures from static values of pressure, mass density, temperature, charge density, number density during oscillations, i.e.

$$\bar{p}_r - \bar{p}_{r,0},\ \bar{\alpha}_r - \bar{\alpha}_{r,0},\ \bar{T}_r - \bar{T}_{r,0},\ \bar{\rho}_r - \rho_{r,0},\ \bar{n}_r - \bar{n}_{r,0}$$

$$\bar{\alpha}_r \simeq \bar{\alpha}_{w,0} \simeq \bar{\alpha}_{n,0} \quad \text{for} \quad r = n, w;$$

$H_r \simeq \bar{H}_{r,0} = \dfrac{R_r \bar{T}_{r,0}}{g}$ = scale height for gas component r (function of z only);

$H \cong H_w \cong H_n$;

$T_0 \cong \bar{T}_{w,0} \cong \bar{T}_{n,0} \cong \bar{T}_{i,0} \cong \bar{T}_{e,0}$, where thermodynamic equilibrium is assumed;

$f_\nu = \dfrac{\sigma\nu}{2\omega}$;

$\bar{\Omega}(\theta,\phi,z,t)$ = tide-producing potential;

k = Boltzmann's constant;

$(gC_v/R)\bar{Q}_r$ = quantity of heat received per unit time and mass by gas component r; assume $\bar{Q} = \bar{Q}_w = \bar{Q}_n$;

$\bar{q}(\theta,\phi,z,t)$ = rate of ion production;

e_r, m_r = atomic or molecular charge and mass for gas r;

$e = e_i = -e_e$;

$\bar{\rho}_e \cong -\bar{\rho}_i$;

D_{rs} = coefficient of mutual diffusion of the two constituents r and s;

$\mathbf{\bar{E}}$ = vector electric field strength;

$\mathbf{B}$ = vector strength of earth's permanent magnetic field;

$\bar{\mathbf{c}}_r$ = mean vector velocity for particles of gas component r where, just as for scalar quantities:

$$\bar{\mathbf{c}}_r(\theta,\phi,z,t) = \sum_{\nu=0}^{\infty} \bar{\mathbf{c}}_{r,\nu}(\theta,\phi,z,t) = \sum_{\nu=0}^{\infty} [\bar{\mathbf{c}}_{r,\nu}(\theta,\phi,z)\, e^{i\sigma_\nu t} + \mathbf{c}_{r,\nu}{}^*(\theta,\phi,z)\, e^{-i\sigma_\nu t}] \tag{2}$$

$\mathbf{\bar{C}}_r$ = mean peculiar velocity for particles of gas component r where:

$$\bar{\mathbf{c}}_r = \bar{\mathbf{c}}_w + \mathbf{\bar{C}}_r$$

$\bar{\mathbf{c}}_w$ = mass-velocity of the gas at any point (Chapman and Cowling, 1952):

$$= \frac{1}{\bar{\alpha}_w} \sum_r \bar{\alpha}_r \bar{\mathbf{c}}_r, \quad \text{where}$$

$$\bar{\alpha}_w = \bar{\alpha}_n + \bar{\alpha}_i + \bar{\alpha}_e \tag{3}$$

$\bar{\mathbf{j}}$ = charge flux-vector (CHAPMAN and COWLING, 1952);

$\bar{\mathbf{j}}_{\text{cond}} = \sum_r n_r e_r \bar{\mathbf{C}}_r$ = conduction current (CHAPMAN and COWLING, 1952); (4)

$\mathbf{F}_r, \mathbf{a}_r$ = applied force per unit mass and acceleration of molecules of gas component r;

$a_\nu = a_\nu(z)$ = attachment coefficient ($\nu = 1$) or recombination coefficient ($\nu = 2$) in the equation for continuity of electrons.

1.2 *Basic equations of motion and energy*

The basic equations of motion and energy for a multi-component gas must be extended as explained in Section 1 to include frictional or collisional effects. The extension for the equations of motion has been performed by JOHNSON [(1951); see footnote (*b*) for a private communication from COWLING] and COWLING (1956). JOHNSON supposes that collisions in a multi-component mixture may be idealized as equivalent to all possible binary collision combinations, and then phenomenologically defines diffusion as

$$\bar{n}_r \bar{\mathbf{c}}_r = -D_{rs} \nabla \bar{n}_r$$

where $D_{rs} = D_{sr}$. His discussion is equivalent to the first diffusion approximation of kinetic theory, that is, to the neglect of thermal diffusion. The general force equation for gas r can be written (JOHNSON, 1951):

$$-\nabla \bar{p}_r + \bar{\alpha}_r(\mathbf{F}_r - \mathbf{a}_r) + \bar{n}_r e_r \bar{\mathbf{c}}_r \times \mathbf{B} = \sum_s \bar{\theta}_{rs}(\bar{\mathbf{c}}_r - \bar{\mathbf{c}}_s)$$

where

$$\bar{\theta}_{rs} = \frac{kT_0}{D_{rs}} \bar{n}_r \bar{n}_s (\bar{n}_r + \bar{n}_s)^{-1}$$

The term $-\bar{\alpha}_r \mathbf{a}_r$ is included according to d'Alembert's principle, where the subscript on $\mathbf{a}$ is a correction to JOHNSON's equation (6). Writing $\mathbf{F}_r$ specifically in terms of electrical and gravitational forces, we obtain the equations of "tidal" motion for a rotating atmosphere in hydrostatic equilibrium

$$\bar{\alpha}_r(i\sigma \bar{\mathbf{c}}_r) + 2\bar{\alpha}_r \boldsymbol{\omega} \times \bar{\mathbf{c}}_w + 2\bar{\alpha}_r \boldsymbol{\omega} \times \bar{\mathbf{C}}_r = -\bar{\alpha}_r \nabla \Omega - \nabla p_r + \bar{\alpha}_r \mathbf{g} + \bar{\rho}_r(\mathbf{E} + \bar{\mathbf{c}}_r \times \mathbf{B}) - \sum_s \bar{\theta}_{rs}(\bar{\mathbf{c}}_r - \bar{\mathbf{c}}_s) \quad \ldots (r = n, e, i) \tag{5}$$

For the charged particles, the equations of motion simplify into an equation for the electrical terms only:

$$\bar{\rho}_r(\mathbf{E} + \bar{\mathbf{c}}_r \times \mathbf{B}) - \sum_s \bar{\theta}_{rs}(\bar{\mathbf{c}}_r - \bar{\mathbf{c}}_s) = 0 \quad \ldots (r = e, i) \tag{6}$$

The last term can be compared to an IR voltage drop and the entire equation to Kirchhoff's voltage law for electrical circuits.

Justification for the reduction to (6) can be found in Appendix 1 where comparison of electrical and mechanical terms is undertaken in an order-of-magnitude treatment; for example, one examines γ_r formed as follows:

$$\gamma_r = \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{\nabla p_r} \right| \simeq \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{p_r / H} \right| \gg \left| \frac{\bar{\mathbf{c}}_r \times \mathbf{B} \bar{\rho}_r H}{\bar{p}_r} \right|$$

Thus the conclusions of Section 1 can be reached, that tidal effects may be considered as occurring in the neutral gas alone and this in turn generates the current through

VOL. 17 1959/6

differential collision of neutrals with charged particles. It is interesting for comparison to note that a Kirchhoff-like voltage law was asserted in the Schuster–Chapman treatment in terms of velocity and current potential functions.

Collisional or viscous terms in the equation of conservation of energy can be included as follows. We will assume that the collisional terms $\sum_s \bar{\theta}_{rs}(\bar{\mathbf{c}}_r - \bar{\mathbf{c}}_s)$ are independent, not of the mean drift velocities, but of the individual particle velocities, and following CHAPMAN and COWLING (1952, Ch. 18) and WILKES (1951):

$$(gC_v\bar{Q}/R)\delta t = \bar{p}\delta(1/\bar{\alpha}) + C_v\,\delta\bar{T} + (\bar{\alpha}_e\mathbf{F}_e.\bar{\mathbf{C}}_e + \bar{\alpha}_i\mathbf{F}_i.\bar{\mathbf{C}}_i + \bar{\alpha}_n\mathbf{F}_n.\bar{\mathbf{C}}_n)\frac{\delta t}{\bar{\alpha}}$$

where $\bar{Q}$, the Wilkes heat source function, may include radiation, convection, and/or conduction effects, and where now

$$\mathbf{F}_r = \frac{e_r\bar{\mathbf{E}}}{m_r} + \frac{e_r}{m_r}\bar{\mathbf{c}}_r \times \mathbf{B} - \frac{1}{\bar{\alpha}_r}\sum_s \bar{\theta}_{rs}(\bar{\mathbf{c}}_r - \bar{\mathbf{c}}_s)$$

If it is assumed that the plasma will always remain nearly neutral ($\bar{n}_e \cong \bar{n}_i$), then

$$\bar{\alpha}_w(gC_v\bar{Q}/R) = \bar{\alpha}_w\bar{p}_w\frac{\delta(1/\bar{\alpha}_w)}{\delta t} + \bar{\alpha}_w C_v\frac{\delta\bar{T}_w}{\delta t} + \bar{\mathbf{j}}_{\text{cond}}.(\bar{\mathbf{E}} + \bar{\mathbf{c}}_w \times \mathbf{B}) - \bar{\theta}_{en}(\bar{\mathbf{C}}_e - \bar{\mathbf{C}}_n)^2 - \bar{\theta}_{in}(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}_n)^2 - \bar{\theta}_{ie}(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}_e)^2 \qquad (7)$$

where (JOHNSON, 1951) $\qquad D_{rs} = D_{sr} \quad \text{or} \quad \bar{\theta}_{rs} = \bar{\theta}_{sr}$

for the gas-as-a-whole (identifying $\bar{\alpha}$ with $\bar{\alpha}_w$, $\bar{p}$ with $\bar{p}_w$ and $\bar{T}$ with $\bar{T}_w$ to conform with present notation). The last three terms are roughly I^2R terms where the $\bar{\theta}_{rn}/\bar{\alpha}_r$ and $\bar{\theta}_{ie}/\bar{\alpha}_e$ factors correspond to collision frequencies or resistances R (see Appendix 2). The signs of the last three terms indicate that the I^2R heating will augment the source function Q and so behave as a possible source function itself. We shall see later how this new source can produce higher harmonic oscillations through non-linear effects.

Before closing this section, we consider the dynamical viscosity terms in the energy conservation equations, that is, those terms arising from collisions of neutrals with neutrals. Employing the tidal velocity expressions, u, v and w, and following MILLER (1951), it can be shown that the mechanical power delivered $-\bar{p}d(\bar{\alpha}^{-1})/dt$ is diminished by the presence of viscous forces

$$-\bar{p}\frac{d(\bar{\alpha}^{-1})}{\alpha t} + \mu\begin{bmatrix}\text{sum of squares of terms involving}\\ \text{derivatives of velocity components}\end{bmatrix}$$

μ is the coefficient of molecular viscosity. The above expression simplifies, for an order of magnitude treatment, to

$$-\bar{p}\frac{d(\bar{\alpha}^{-1})}{dt} + \frac{4}{3}\bar{\alpha}^{-1}\mu\left(\frac{\partial w}{\partial z}\right)^2 + \bar{\alpha}^{-1}\mu\left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2\right] \qquad (8)$$

where terms involving $\partial/a\partial\theta$ and $\partial/a\sin\theta\,\partial\phi$ are less than the $\partial/\partial z$ terms by a factor $\sim H/a$ and have hence been neglected; the needed azimuthal dependence for u, v and w are given by their linear expressions (WHITE, 1956).

2. FUNDAMENTAL EQUATIONS FOR RESONANCE THEORY

The basic equation of energy (7) for the gas-as-a-whole may be simplified by employing equations (6) and by noting $\bar{\theta}_{rs} = \bar{\theta}_{sr}$; thus

$$\left.\begin{aligned} &\bar{\alpha}_w(gC_v\bar{J}_w/R_w) = \bar{\alpha}_w\bar{p}_w\,\delta(\bar{\alpha}_w{}^{-1})/\delta t + \bar{\alpha}_w C_v\,\delta\bar{T}_w/\delta t \\ &\text{where } \bar{J}_w \text{ and } \bar{q}_w \text{ are defined by} \\ &\bar{J}_w = \bar{Q} + \bar{q}_w = \bar{Q} + \frac{R_w}{gC_v\bar{\alpha}_w}\,[\bar{\theta}_{ni}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_i) + \bar{\theta}_{ne}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_e)]\,.\,\bar{\mathbf{C}}_n \end{aligned}\right\} \tag{9}$$

Following the procedure already used in obtaining equation (7), the first law of thermodynamics for the neutral component becomes

$$\bar{\alpha}_n(gC_v\bar{Q}/R) = \bar{\alpha}_n\bar{p}_n\,\delta(\bar{\alpha}_n{}^{-1})/\delta t + \bar{\alpha}_n C_v\,\delta\bar{T}_n/\delta t - \bar{\theta}_{ni}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_i)^2 - \bar{\theta}_{ne}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_e)^2 \tag{10}$$

where the "solar" source function $\bar{Q}$ is, for convenience, not distinguished in equations (9) and (10). (Physically, this implies that all thermal energy which induces oscillations is absorbed by the neutral gas component alone.) Setting

$$\left.\begin{aligned} &\bar{J}_n = \bar{Q} + \bar{q}_n = \bar{Q} + \frac{R_n}{gC_v\bar{\alpha}_n}\,[\bar{Q}_{ni}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_i)^2 + \bar{\theta}_{ne}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_e)^2] \\ &\text{then} \qquad \bar{\alpha}_n(gC_v\bar{J}_n/R_n) = \bar{\alpha}_n\bar{p}_n\,\delta(\bar{\alpha}_n{}^{-1})/\delta t + \bar{\alpha}_n C_v\,\delta\bar{T}_n/\delta t \end{aligned}\right\} \tag{11}$$

where $\bar{q}_n > 0$. $\bar{q}_n$ and $\bar{q}_w$ are related by

$$\bar{q}_n \cong \bar{q}_w + \frac{R}{gC_v\bar{\alpha}_n}\,[\bar{\theta}_{ni}(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}_n)\,.\,\bar{\mathbf{C}}_i + \bar{\theta}_{ne}(\bar{\mathbf{C}}_e - \bar{\mathbf{C}}_n)\,.\,\bar{\mathbf{C}}_e] \tag{12}$$

Comparing equations (9) and (11) (where $R = R_w = R_n$) with the basic energy conservation equation of WILKES (1951), we see that, in form, the first law of thermodynamics remains unchanged. In like manner we define new potential functions

$$\bar{\Lambda}_r = \bar{\Omega} + \bar{\pi}_r \quad \ldots (r = w, n) \tag{13}$$

so that collisional effects in the equations of motion are simply incorporated into the potential function. From equations (5) and (6) we find that if

$$\bar{\boldsymbol{\Pi}}_r = \nabla\bar{\pi}_r \tag{14}$$

where $\bar{\boldsymbol{\Pi}}_r$ is defined by

$$\bar{\boldsymbol{\Pi}}_r = \bar{\alpha}_r{}^{-1}[\bar{\theta}_{ne}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_e) + \bar{\theta}_{ni}(\bar{\mathbf{C}}_n - \bar{\mathbf{C}}_i)] \quad \ldots (r = w, n)$$

or

$$\bar{\boldsymbol{\Pi}} \cong \bar{\boldsymbol{\Pi}}_w \cong \bar{\boldsymbol{\Pi}}_n \tag{15}$$

then the equations of motion take the same form as given by WILKES (1954),

$$\bar{\alpha}_r\frac{\partial\bar{\mathbf{c}}_r}{\partial t} + 2\bar{\alpha}_r\boldsymbol{\omega}\times\bar{\mathbf{c}}_r = -\bar{\alpha}_r\nabla\Lambda_r - \nabla p_r + \bar{\alpha}_r\mathbf{g} \quad \ldots (r = w, n) \tag{16}$$

Regarding the other basic tidal equations, we have for static equilibrium,

$$\frac{d\bar{p}_{r,0}}{dz} + g\bar{\alpha}_{r,0} = 0 \quad \ldots (r = w, n)$$

for the perfect gas law, $\qquad \bar{p}_r = R_r\bar{\alpha}_r\bar{T} \quad \ldots (r = w, n)$

VOL. 17 1959/6

assuming thermodynamic equilibrium, and for the equation of continuity, neglecting sources and sinks for the present discussion,

$$\frac{D_r \bar{\alpha}_r}{D_r t} + \bar{\alpha}_r \chi_r = 0 \quad \ldots (r = w, n)$$

where

$$\bar{\chi}_r = \nabla . \bar{\mathbf{c}}_r \quad \ldots (r = w, n) \tag{17}$$

and where

$$\frac{D_r}{D_r t} = \frac{\partial}{\partial t} + \bar{\mathbf{c}}_r . \nabla \quad \ldots (r = w, n)$$

It is thus seen that all the basic "tidal" equations retain the form of the "tidal" treatment; the results of previous oscillatory theory should therefore follow (*mutatis mutandis*). We must, however, depart somewhat from previous treatment. The presence of non-linear terms $\bar{q}_r$ and $\bar{\pi}_r$ makes the use of an assumed time factor $e^{i\sigma}$ here inapplicable, since, for example,

$$\cos^2 \sigma_n t \neq \text{real part } e^{2i\sigma_n t}$$

One is therefore obliged to use real functions. Thus, for example, representation (1) (HINES, 1953) of a Fourier expansion for the time-dependent quantities will be used; time derivatives (and also integrations) of p_r will retain the same form thus

$$\frac{\partial \bar{p}_r}{\partial t} = \sum_{\nu=0}^{\infty} (i\sigma_\nu \bar{p}_{r,\nu}) = \sum_{\nu=0}^{\infty} [(i\sigma_\nu p_{r,\nu}) e^{i\sigma_\nu t} + (i\sigma_\nu p_r)^* e^{-i\sigma_\nu t}]$$

and

$$\int \bar{p}_r \, dt = \sum_{\nu=0}^{\infty} \left(\frac{\bar{p}_{r,\nu}}{i\sigma_\nu} \right) = \sum_{\nu=0}^{\infty} \left[\left(\frac{p_{r,\nu}}{i\sigma_\nu} \right) e^{i\sigma_\nu t} + \left(\frac{p_{r,\nu}}{i\sigma_\nu} \right)^* e^{-i\sigma_\nu t} \right] \tag{18}$$

Second degree non-linearities such as $\nabla \bar{\pi}_r$ may be quickly linearized as follows (HINES, 1953):

$$4\bar{A}\bar{B} = \sum_{\eta=1}^{\infty} \left\{ \left(\sum_{r=0}^{\eta} A_r B_{\eta - r} + \sum_{r=0}^{\infty} A_r^* B_{\eta + r} + \sum_{r=\eta}^{\infty} A_r B_{r-\eta}^* \right) e^{i\sigma_\eta t} \right.$$
$$\left. + \sum_{\eta=1}^{\infty} \{(\ldots)^* e^{-i\sigma_\eta t}\} + A_0 B_0 + A_0^* B_0^* + \sum_{\eta=0}^{\infty} \{A_\eta^* B_\eta + A_\eta B_\eta^*\} \right\} \tag{19}$$

where A and B are arbitrary functions. For $\bar{q}_w$, where the non-linearity is of the third degree, linearization in time can be achieved by simple reapplication of the above equation [see equation (42)].

The assumption involved in the outlined procedures are (1) a time factor exists and (2) the time factor may be expanded as a Fourier series.

Once the basic tidal equations have been linearized in time, each equation may be replaced by an infinite number of equations, that is, one for each harmonic $e^{i\sigma_n t}$ (or $e^{-i\sigma_n t}$) and the set of equations (9), (16) and (17) may then be replaced by any infinite set, one for each harmonic $e^{i\sigma_n t}$ (and a complex conjugate set for $e^{-i\sigma_n t}$). Applying the tidal treatment to each set, one can, for example, recover the "tidal" equations (WILKES, 1949) for $u_{r,\nu}$ and $v_{r,\nu}$

$$u_{r,\nu} = \frac{\sigma_\nu}{4a\omega^2 (f_\nu^2 - \cos^2 \theta)} \left[i \frac{\partial}{\partial \theta} + \frac{\cos \theta}{f_\nu} \frac{\partial}{\partial \phi} \right] \left(\frac{\bar{p}_{r,\nu}}{\bar{\alpha}_{r,0}} + \bar{\Lambda}_{r,\nu} \right) \tag{20}$$

$$\bar{v}_{r,\nu} = \frac{i\sigma_\nu}{4a\omega^2 (f_\nu^2 - \cos^2 \theta)} \left[\frac{i \cos \theta}{f_\nu} \frac{\partial}{\partial \theta} + \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right] \left(\frac{\bar{p}_{r,\nu}}{\bar{\alpha}_{r,0}} + \bar{\Lambda}_{r,\nu} \right) \tag{21}$$

where $\bar{p}_{r,\nu}$ and $i\sigma_\nu \bar{p}_{r,\nu}$ have the representations given in equations (1) and (18), and similarly for $\bar{\Lambda}_{r,\nu}$ and $i\sigma_\nu \bar{\Lambda}_{r,\nu}$. The time-dependent wave-equation (WILKES, 1951) valid for both thermal and gravitational excitation (since the gravitational potential appears only implicitly) also retains the same form,

$$\gamma H_r \frac{\partial^2 \bar{\chi}_{r,\nu}}{\partial z^2} + \gamma\left(\frac{dH_r}{dz} - 1\right)\frac{\partial \bar{\chi}_{r,\nu}}{\partial z} + \frac{\partial}{\partial z}\left[\frac{1}{H_r}\left(\frac{dH_r}{dz} + 1\right)J_{r,\nu} - \frac{\partial J_{r,\nu}}{\partial z}\right]$$
$$+ \frac{g}{4a^2\omega^2} F_\nu\left\{\left(1 - \gamma - \gamma\frac{dH}{dz}\right)\bar{\chi}_{r,\nu} + \frac{1}{H_r}\left(\frac{dH_r}{dz} + 1\right)J_{r,\nu}\right\} \cong 0 \quad \ldots (r = w, n) \quad (22)$$

where the differential operator

$$F_\nu = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\frac{\sin\theta}{f_\nu^2 - \cos^2\theta}\left(\frac{\partial}{\partial\theta} - \frac{i\cot\theta}{f_\nu}\frac{\partial}{\partial\phi}\right)\right] + \frac{1}{f_\nu^2 - \cos^2\theta}\left[\frac{i\cot\theta}{f_\nu}\frac{\partial^2}{\partial\theta\,\partial\phi} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$$

where it is assumed that $\partial^2\bar{\Lambda}_{r,\nu}/\partial z^2 = 0$ and where the dependent variables $\bar{\chi}_{r,\nu}$ and $J_{r,\nu}$ assume the form given in equations (1) for the pressure. Use of the subscript ν points up the modal treatment in both linear and non-linear theory. In the following section, a new significance will be attached to f_ν, namely the role of a differential operator.

Use of a potential function [equation (14)] involves a rather serious assumption. It is quite unlikely that terms depending on collisional effects can be described as the gradient of a scalar. The prime advantage of such an assumption is that the total wave-equation (22) is soluble by the method of separation of variables. However, a variable separable solution may still be achieved by replacing equation (15) by the conditions that

$$\bar{\Pi}_\theta = \frac{\partial\bar{\pi}}{a\partial\theta} \quad \text{and} \quad \bar{\Pi}_\phi = \frac{\partial\bar{\pi}}{a\sin\theta\,\partial\phi}$$

but $\bar{\Pi}_z$ is not to be expressed as the gradient of a scalar potential. If collision frequencies and temperatures or scale heights may be considered essentially functions of z alone, then one has overcome the difficulty of expressing $\bar{\theta}_{rs}$ as the gradient of a scalar. Further, the differential wave-equation will continue to be variables separable; with the above substitution for $\bar{\Pi}_\theta$ and $\bar{\Pi}_\phi$, the equations for u and v are still given by equations (20) and (21), respectively, while the differential wave-equation now becomes

$$\gamma H_r \frac{\partial^2 \bar{\chi}_{r,\nu}}{\partial z^2} + \gamma\left(\frac{dH_r}{dz} - 1\right)\frac{\partial \bar{\chi}_{r,\nu}}{\partial z} + \frac{\partial}{\partial z}\left[\frac{1}{H_r}\left(1 + \frac{dH_r}{dz}\right)J_{r,\nu} - \frac{\partial J_{r,\nu}}{\partial z} - \frac{i\sigma_\nu}{g}\bar{\Pi}_{z,\nu}\right]$$
$$+ \frac{g}{4a^2\omega^2} F_\nu\left\{\left(1 - \gamma - \gamma\frac{dH_r}{dz}\right)\bar{\chi}_{r,\nu} + \frac{1}{H_r}\left(\frac{dH_r}{dz} + 1\right)J_{r,\nu}\right.$$
$$\left. - \frac{i\sigma_\nu}{g}\left(\bar{\Pi}_{z,\nu} - \frac{\partial\bar{\pi}_\nu}{\partial z}\right)\right\} = 0 \quad \ldots (r = w, n) \quad (23)$$

2.1 *The wave-equation including the time*

Considerable simplification of wave-equations (22) obtains if one lets

$$\bar{\chi}_{r,\nu} = \frac{J_{r,\nu}}{\gamma H_r} + \bar{\psi}_{r,\nu}(\theta, \phi, z, t) \quad (24)$$

VOL. 17 1959/6

where $\psi_{r,\nu}$ as defined here shall replace the previous definition of WILKES (1951) for ψ, namely, $\chi(\theta,\phi,z)=\chi(z)\,\psi(\theta,\phi)$. Wave-equation (22) now becomes

$$\gamma H_r\frac{\partial^2\bar{\psi}_{r,\nu}}{\partial z^2}+\gamma\left(\frac{dH_r}{dz}-1\right)\frac{\partial\bar{\psi}_{r,\nu}}{\partial z}+\frac{g}{4a^2\omega^2}F_\nu\left\{\left(1-\gamma-\gamma\frac{dH_r}{dz}\right)\bar{\psi}_{r,\nu}+\frac{J_{r,\nu}}{\gamma H_r}\right\}=0$$
$$\ldots(r=w,n) \tag{25}$$

without any need, thus far, of a variables separable condition in $\psi_{r,\nu}(\theta,\phi,z)$; this is more general than previous treatments (compare with WHITE, 1956). The physical interpretation put on equation (24) is that $\bar{\psi}_{r,\nu}$ represents the resonant solution while $J_{r,\nu}/\gamma H_r$ represents the solution for the forced oscillation. It would be of great interest to determine the relative importance of the two solutions. Nevertheless, this interpretation may tentatively account (Section 5.3) for h-values of certain observed non-linear vibrations being far removed from the resonant value.

We shall now employ the method of operators to obtain a time-dependent wave-equation. Such an equation will have many uses, one immediate use being the establishment of the stability criterion. For the purpose of mathematical simplification, we shall limit our discussion to low and intermediate latitudes ($f_\nu^2\gg\cos^2\theta$), so that the differential operator F_ν reduces to

$$F_\nu\cong\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\frac{\sin\theta}{f_\nu^2}\left(\frac{\partial}{\partial\theta}-\frac{i\cot\theta}{f_\nu}\frac{\partial}{\partial\phi}\right)\right]+\frac{1}{f_\nu^2}\left[\frac{i\cot\theta}{f_\nu}\frac{\partial^2}{\partial\theta\,\partial\phi}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{26}$$

With this approximation to F_ν, equation (25) can immediately be written as a cubic equation in f_ν. Now as an example, considering equations (1) and (18) for the pressure

$$\frac{\partial^\eta\bar{p}_{r,\nu}}{\partial t^\eta}=[(i\sigma_\nu)^\eta p_{r,\nu}]\,e^{i\sigma_\nu t}+[(i\sigma_\nu)^\eta p_{r,\nu}]^*\,e^{-i\sigma_\nu t}=(i\sigma_\nu)^\eta\bar{p}_{r,\nu}=D_t\bar{p}_{r,\nu}$$

where ν and η are positive integer. It is thus possible to replace $(i\sigma_\nu)^\eta$ by an operator in time. Recalling the definition $f_\nu=\sigma_\nu/2\omega$, we have

$$(if_\nu)^\eta\bar{p}_{r,\nu}=\frac{1}{(2\omega)^\eta}D_t^\eta\bar{p}_{r,\nu} \tag{27}$$

so that equation (25), together with (26), becomes a differential equation of the third order in time, namely,

$$\begin{aligned}&\frac{1}{(2i\omega)^3}\left[\gamma H_r\frac{\partial^2}{\partial z^2}+\gamma\left(\frac{dH_r}{dz}-1\right)\frac{\partial}{\partial z}\right]\frac{\partial^3\bar{\psi}_r}{\partial t^3}\\&+\frac{1}{2i\omega}\frac{g}{4a^2\omega^2}\frac{\partial}{\partial\theta}\left[\frac{1}{\sin\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\left[\left(1-\gamma-\gamma\frac{dH_r}{dz}\right)\frac{\partial\bar{\psi}_r}{\partial t}+\frac{1}{\gamma H_r}\frac{\partial J_r}{\partial t}\right]\\&+\frac{g}{4a^2\omega^2}\left[-\frac{i}{\sin\theta}\frac{\partial}{\partial\theta}\left(\cos\theta\frac{\partial}{\partial\phi}\right)+i\cot\theta\frac{\partial^2}{\partial\theta\,\partial\phi}\right]\left[\left(1-\gamma-\gamma\frac{dH}{dz}\right)\bar{\psi}_r+\frac{J_r}{\gamma H_r}\right]=0\\&\qquad\ldots(r=w,n)\end{aligned} \tag{28}$$

Use of the total wave-function $\bar{\psi}_r$ in equation (28) may be justified as follows: It is clear that equation (28) governs the behaviour of the component $\psi_{r,\nu}\,e^{i\sigma_\nu t}$; the equation conjugate to this equation, i.e. i replaced by $-i$, holds for the component $\psi_{r,\nu}{}^*\,e^{-i\sigma_\nu t}$; but equation (28) and its conjugate will take the same form, so that upon addition, the resulting equation will maintain form (28) where now the wave-function satisfying

equation (28) takes the form $\bar{\psi}_{r,\nu} = \psi_{r,\nu} e^{i\sigma_\nu t} + \psi_{r,\nu}^* e^{-i\sigma_\nu t}$. Summing over all harmonics, $\bar{\psi}_r = \sum_{\nu=0}^{\infty} \bar{\psi}_{r,\nu}$ becomes the solution of the total wave-equation (28).

In conclusion, it is seen that within the framework of present approximations, one may still talk of "modes" in the non-linear oscillatory theory.

2.2 *Steady winds*

Before closing this section, it can immediately be seen from the time-dependent wave-equation (28) that one can for the first time directly obtain the linear wave-function corresponding to the steady state winds (seasonal winds), thus, for a particular solution,

$$\bar{\psi}_{r,0}(z) = -\frac{1}{\gamma H_r} \frac{\bar{Q}_{r,0}(z)}{(1-\gamma-\gamma dH_r/dz)} \tag{29}$$

From the discussion of Section 5.1, equation (72) to follow, on the representation of the source function for ionized particles, it follows that equation (29) can be written as

$$\psi_0(z) \simeq -\frac{Q_{0,0}+Q_{1,0}\sin\theta+Q_{2,0}\sin^2\theta}{\gamma H(1-\gamma-\gamma dH/dz)} \tag{30}$$

where the subscript r has either been deleted or replaced by an index for delineating the dependence on latitude; the second subscript still indicates the steady state or the zeroth harmonic. While equations (29) and (30) do not include possible non-linear contributions to the steady winds, these equations should be useful for obtaining "tidal" winds for the first 120 km or so of the earth's atmosphere.

3. Derivation of Expressions for the Drift Velocities $\bar{\mathbf{c}}_e$ and $\bar{\mathbf{c}}_i$, the Non-linear Source Functions $\bar{q}_w$ and $\bar{q}_n$, the Electric Field $\bar{\mathbf{E}}$, and the Current Density $\bar{\mathbf{j}}$

At this point, it will be worthwhile to collect the equations which are important to a final solution of the tidal–dynamo problem. They are wave-equations (25) for the gas-as-a-whole and for the neutral component, Kirchhoff's voltage law (6), for both positive ions and electrons, the equation (3) defining mass velocity, $\bar{\mathbf{c}}_w$; since the electron density which appears in many of the equations is a function of time, it is necessary to consider the problem of layer formation through the electron continuity equation (68). Two vector unknowns in the problem are $\bar{\mathbf{c}}_w$ and $\bar{\mathbf{c}}_n$ found from χ_w and χ_n through equations akin to equations (10), (11) and (12) of WHITE (1956); the other vector unknowns are $\bar{\mathbf{E}}$, $\bar{\mathbf{c}}_e$, and $\bar{\mathbf{c}}_i$; the electron density is a final scalar unknown. We may thus summarize the tidal–dynamo problem as consisting of three non-linear differential equations and three scalar unknowns $\bar{\chi}_w$, $\bar{\chi}_n$, and $\bar{\rho}_e$, plus three vector equations [Kirchhoff's voltage laws and definition (3)] and three vector unknowns $\bar{\mathbf{E}}$, $\bar{\mathbf{c}}_e$ and $\bar{\mathbf{c}}_i$. In the present case, it will be convenient to express $\bar{\mathbf{c}}_i$, $\bar{\mathbf{c}}_e$ and $\bar{\mathbf{E}}$, as well as $\bar{q}_w$, $\bar{q}_n$ and $\bar{\mathbf{j}}$ in terms of $\bar{\mathbf{c}}_w$ and $\bar{\mathbf{c}}_n$. Justification for this statement is as follows: The neutral component of the atmosphere, being by far the largest component, follows the classical "tidal" theory and, in fact, as a good first approximation, $\bar{\chi}_{w,\nu} = \bar{\chi}_{n,\nu}$ and $\bar{\mathbf{c}}_{w,\nu} = \bar{\mathbf{c}}_{n,\nu} = \bar{\mathbf{c}}_{e,\nu} = \bar{\mathbf{c}}_{i,\nu}$. Substituting the values $\bar{\mathbf{c}}_{w,\nu}$ or $\bar{\mathbf{c}}_{n,\nu}$ into equations (40) and (41) for $\bar{q}_{w,\nu} = fct.(\bar{\mathbf{c}}_{w,\nu}; \bar{\mathbf{c}}_{n,\nu})$ and $\bar{q}_{n,\nu} = fct.(\bar{\mathbf{c}}_{w,\nu}; \bar{\mathbf{c}}_{n,\nu})$, respectively, it appears that $\bar{q}_{w,\nu} = 0$ and $\bar{q}_{n,\nu} \neq 0$. Since $\bar{\chi}_{w,\nu}$ and $\bar{\chi}_{n,\nu}$ depend upon $\bar{q}_{w,\nu}$ and $\bar{q}_{n,\nu}$,

VOL. 17 1959/6

respectively, through the differential wave-equation, $\bar{\chi}_{w,\nu} \mp \bar{\chi}_{n,\nu}$ and substitution into the linear expressions for $\bar{\mathbf{c}}_{w,\nu}$ and $\bar{\mathbf{c}}_{n,\nu}$ yields $\bar{\mathbf{c}}_{w,\nu} \mp \bar{\mathbf{c}}_{n,\nu}$. Repeating the process until it converges, one can then determine $\bar{\mathbf{c}}_{i,\nu}$, $\bar{\mathbf{c}}_{e,\nu}$, $\bar{\mathbf{E}}_\nu$, and $\bar{\mathbf{j}}_\nu$ if they are functions of $\bar{\mathbf{c}}_{w,\nu}$ and $\bar{\mathbf{c}}_{n,\nu}$ only.

The method of successive approximation envisaged above is not at all essential to what follows; the only approximations made will be those based in a very general way on the atmospheric model set down in Table 1.

The quantities $\bar{\mathbf{c}}_e$, $\bar{\mathbf{c}}_i$, and $\bar{\mathbf{E}}$ pertaining to the plasma can be found by appealing only to the algebraic electrical circuit laws (6) and the equation defining $\bar{\mathbf{c}}_w$. Since $\bar{\mathbf{c}}_w$ and $\bar{\mathbf{c}}_n$ can be considered as knowns, we have the same number of vector equations as vector unknowns. In the numerous treatments of the dynamo problem, there have been more unknowns than equations and the problem is met by employing idealized boundaries which place somewhat artificial restrictions† on $\bar{\mathbf{E}}$. In the present discussion, there should be no need to introduce artificially stratified layers or characteristic lengths which give the scale of the current systems. The latter should follow naturally from the micro- and macroscopic parameters of the problem.

Expressions for the drift velocities $\bar{\mathbf{c}}_r$ may be obtained as follows. From Kirchhoff's voltage laws (6), the electric field $\bar{\mathbf{E}}$ may be eliminated thus:

$$\bar{\rho}_e(\bar{\mathbf{c}}_e - \bar{\mathbf{c}}_i) \times \mathbf{B} - \bar{\theta}_{en}\,\bar{\mathbf{c}}_e - \bar{\theta}_{in}\,\bar{\mathbf{c}}_i + (\bar{\theta}_{en} + \bar{\theta}_{in})\,\bar{\mathbf{c}}_n = 0. \tag{31}$$

From equations (3) and (31), $\bar{\mathbf{c}}_i$ may now be eliminated, obtaining

$$\bar{\mathbf{c}}_e - \beta_{en}\,\Gamma^{-1}\,\bar{\mathbf{c}}_e \times \mathbf{B} = \Gamma^{-1}\left[\Gamma\bar{\mathbf{c}}_n + \frac{\bar{\alpha}_w \beta_{en}}{\bar{\alpha}_i \beta_{in}}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) - \frac{\bar{\alpha}_w \beta_{en}}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) \times \mathbf{B} - \beta_{en}\,\bar{\mathbf{c}}_n \times \mathbf{B}\right] \tag{32}$$

where we define

$$\beta_{en} = \bar{\rho}_e/\bar{\theta}_{en}, \quad \beta_{in} = \bar{\rho}_i/\bar{\theta}_{in}, \quad \text{and} \quad \Gamma = 1 + \bar{\alpha}_e \beta_{en}/\bar{\alpha}_i \beta_{in}$$

and where the factor $(1 + \bar{\alpha}_e/\bar{\alpha}_i)$ has been replaced by unity. Solving for $\bar{\mathbf{c}}_e$,

$$\bar{\mathbf{c}}_e = \left(\frac{\beta_{en}}{\beta_{in}}\Gamma^{-1} - 1\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_\parallel + \left(\frac{\Gamma^2}{{\beta_{en}}^2 B^2} + 1\right)\bar{\mathbf{c}}_n + \left(1 + \frac{\Gamma}{\beta_{en}\beta_{in}B^2}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) + \frac{1}{B^2}\left(\frac{\Gamma}{\beta_{en}} - \frac{1}{\beta_{in}}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B} \times (\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) \tag{33}$$

where the subscript $\parallel$ indicates the component of the vector parallel to the magnetic field $\mathbf{B}$ and where the following very close approximation (see Appendix 2) has been introduced,

$$1 + {\beta_{en}}^2 \Gamma^{-2} B^2 = {\beta_{en}}^2 \Gamma^{-2} B^2 \tag{34}$$

In like manner, the equation for $\bar{\mathbf{c}}_i$ in terms of the knowns $\bar{\mathbf{c}}_w$ and $\bar{\mathbf{c}}_n$ may be obtained,

$$\bar{\mathbf{c}}_i = (\Gamma^{-1} - 1)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_\parallel + \left(\frac{\Gamma^2}{{\beta_{en}}^2 B^2} + 1\right)\bar{\mathbf{c}}_n + \left(\frac{\Gamma}{{\beta_{en}}^2 B^2} + 1\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) + \frac{\Gamma - 1}{\beta_{en} B^2}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B} \times (\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) \tag{35}$$

† These restrictions, as well as those enumerated in the Introduction, are prevalent in the many articles devoted wholly or in part to dynamo theory as in the *Journal of Geomagnetism and Geoelectricity*, from Vol. 1, 1949, to the present.

By consideration of the physical conditions of the problem according to Appendix 2, equations (33) and (35) may be somewhat simplified, thus

$$\bar{\mathbf{c}}_e - \frac{\beta_{en}\bar{\alpha}_w}{\Gamma\beta_{in}\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\parallel} + \bar{\mathbf{c}}_n + \left(1 + \frac{\Gamma}{\beta_{en}\beta_{in}B^2}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) - \frac{1}{\beta_{in}B^2}\mathbf{B}\times(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) \tag{36}$$

and

$$\bar{\mathbf{c}}_i \simeq -\frac{\beta_{en}\bar{\alpha}_w\bar{\alpha}_e}{\Gamma\beta_{in}\bar{\alpha}_i^{\,2}}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\parallel} + \bar{\mathbf{c}}_n + \frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) + \frac{1}{\beta_{in}B^2}\frac{\bar{\alpha}_e\bar{\alpha}_w}{\bar{\alpha}_i^{\,2}}\mathbf{B}\times(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n) \tag{37}$$

From definition (15) and equations (33) and (35),

$$\begin{aligned}\bar{\boldsymbol{\Pi}} = -\frac{1}{\bar{\alpha}_i}\Bigg\{&\left[\left(\frac{1-\Gamma}{\Gamma}\right)\bar{\theta}_{ni} + \left(\frac{\beta_{en}}{\Gamma\beta_{in}} - 1\right)\bar{\theta}_{ne}\right](\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\parallel}\\ &+\left[\left(1 + \frac{\Gamma}{\beta_{en}^{\,2}B^2}\right)\bar{\theta}_{ni} + \left(1 + \frac{\Gamma}{\beta_{en}\beta_{in}B^2}\right)\bar{\theta}_{ne}\right](\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)\\ &+\left[\left(\frac{\Gamma-1}{\beta_{en}}\right)\bar{\theta}_{ni} + \left(\frac{\Gamma}{\beta_{en}} - \frac{1}{\beta_{in}}\right)\bar{\theta}_{ne}\right]\frac{1}{B^2}\mathbf{B}\times(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)\Bigg\}\end{aligned} \tag{38}$$

or

$$\bar{\boldsymbol{\Pi}} \simeq -\frac{1}{\bar{\alpha}_i}\left[\bar{\theta}_{in}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\perp} - \frac{\bar{\theta}_{en}}{\beta_{in}B^2}\mathbf{B}\times(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)\right] \tag{39}$$

where the subscript $\perp$ denotes the component of the vector perpendicular to $\mathbf{B}$.

To obtain $\bar{q}_w$, substitute equations (33) and (35) for $\bar{\mathbf{c}}_e$ and $\bar{\mathbf{c}}_i$, respectively, into equation (9) for $\bar{q}_w$, obtaining

$$\bar{q}_w = \frac{R\bar{\rho}_e}{gC_v\bar{\alpha}_w}\left(\frac{1}{\beta_{en}} - \frac{1}{\beta_{in}}\right)\left[\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\perp}^{\,2} + \frac{1}{\beta_{en}^{\,2}B^2}\Gamma^2(\bar{\mathbf{c}}_w\cdot\bar{\mathbf{c}}_n - \bar{\mathbf{c}}_n^{\,2})\right] \tag{40}$$

The expression for $\bar{q}_n$ may be obtained in a straightforward manner from definition (11), and equations (33) and (35). Unfortunately, the exact expression for $\bar{q}_n$ is quite cumbersome. Some simplification obtains through appealing to the physical conditions of the problem (Appendix 2), resulting in an approximate expression

$$\begin{aligned}\bar{q}_n \simeq \left(\frac{R\bar{\rho}_e}{gC_v\bar{\alpha}_w}\right)\Bigg\{&\Gamma^{-2}\frac{\beta_{en}}{\beta_{in}^{\,2}}\frac{\bar{\alpha}_w^{\,2}}{\bar{\alpha}_i^{\,2}}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\parallel}^{\,2} - \frac{\Gamma^4}{\beta_{in}\beta_{en}^{\,4}B^4}\bar{\mathbf{c}}_n^{\,2}\\ &-\frac{2\Gamma^2}{\beta_{in}\beta_{en}^{\,2}B^2}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\bar{\mathbf{c}}_n\cdot(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)_{\perp} + \frac{2\Gamma^3}{\beta_{in}\beta_{en}^{\,3}B^4}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\bar{\mathbf{c}}_n\cdot(\bar{\mathbf{c}}_w\times\mathbf{B})\\ &-\frac{1}{\beta_{in}}\frac{\bar{\alpha}_w^{\,2}}{\bar{\alpha}_i^{\,2}}(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)^2 + \frac{1}{\beta_{en}\beta_{in}^{\,2}B^4}\frac{\bar{\alpha}_w^{\,2}}{\bar{\alpha}_i^{\,2}}[\mathbf{B}\times(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)]^2\Bigg\}\end{aligned} \tag{41}$$

which should be accurate between 100 and 275 km and higher. $\bar{q}_r$ can be represented by a doubly infinite series; this representation will be useful for Section 4. From equation (19),

$$\begin{aligned}(4\bar{A}\bar{B})(2\bar{C}) = \sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty}&[C_\nu(\ldots)]\,e^{i\sigma_\nu t}e^{i\sigma_\eta t} + [C_\nu(\ldots)]^*\,e^{-i\sigma_\nu t}e^{-i\sigma_\eta t}\\ &+[C_\nu{}^*(\ldots)]\,e^{-i\sigma_\nu t}e^{i\sigma_\eta t} + [C_\nu{}^*(\ldots)]^*\,e^{i\sigma_\nu t}e^{-i\sigma_\eta t}\end{aligned} \tag{42}$$

VOL. 17 1959/6

where $\bar{C}$ is also an arbitrary function and where

$$(\ldots) = \left(\sum_{r=0}^{\eta} A_\eta B_{\eta-r} + \sum_{r=0}^{\infty} A_r{}^* B_{\eta+r} + \sum_{r=\eta}^{\infty} A_r B_{r-\eta}{}^* \right)$$

$\bar{C}$ here might correspond to $\bar{\rho}_e$ and $\bar{A}\bar{B}$ to $(\bar{\mathbf{c}}_w - \bar{\mathbf{c}}_n)^2$, say. Thus $\bar{q}_r$ may be reduced from the third to the second degree in time.

Now that expressions for $\bar{\mathbf{c}}_e$ and $\bar{\mathbf{c}}_i$ are given, an expression for $\bar{\mathbf{E}}$ can be found from equation (6) with $r = e$, thus

$$\begin{aligned}\bar{\mathbf{E}} = {} & \left[\Gamma^{-1}\frac{\beta_{en}}{\beta_{in}}\left(\frac{1}{\beta_{ei}}+\frac{1}{\beta_{en}}\right)-\frac{1}{\beta_{en}}-\frac{\Gamma^{-1}}{\beta_{ei}}-\frac{1}{\beta_{in}}+\frac{\Gamma}{\beta_{en}}\right]\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)_\parallel \\ & + \frac{\Gamma^2}{\beta_{en}{}^3 B^2}\bar{\mathbf{c}}_n + \left[\left(\frac{1}{\beta_{ei}}+\frac{1}{\beta_{en}}\right)\left(\frac{1}{\beta_{in}\beta_{en}B^2}\right)\Gamma - \frac{\Gamma}{\beta_{ei}\beta_{en}{}^2 B^2}+\frac{1}{\beta_{in}}-\frac{\bar{\alpha}_e}{\beta_{in}\bar{\alpha}_i}\right]\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) \\ & + \left[\left(\frac{1}{\beta_{ei}}+\frac{1}{\beta_{en}}\right)\frac{\Gamma}{\beta_{en}B^2}-\frac{1}{\beta_{ei}\beta_{in}B^2}-\frac{\bar{\alpha}_e}{\bar{\alpha}_i\beta_{ei}\beta_{in}B^2}+1+\frac{\bar{\alpha}_e}{\bar{\alpha}_i\beta_{in}{}^2 B^2}\right]\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B}\times(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) \\ & + \left[\frac{\Gamma^2}{\beta_{en}{}^2 B^2}+1\right]\mathbf{B}\times\bar{\mathbf{c}}_n \end{aligned} \tag{43}$$

where

$$\beta_{ei} = \bar{\rho}_e/\bar{\theta}_{ei}$$

or employing again the values given in Appendix 2 for actual ionospheric condition,

$$\begin{aligned}\bar{\mathbf{E}} \simeq {} & \left[\Gamma^{-1}\frac{\beta_{en}}{\beta_{in}}\left(\frac{1}{\beta_{ei}}+\frac{1}{\beta_{en}}\right)-\frac{1}{\beta_{in}}\right]\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)_\parallel + \frac{\Gamma^2}{\beta_{en}{}^3 B^2}\bar{\mathbf{c}}_n + \frac{1}{\beta_{in}}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) \\ & + \left[1+\frac{\bar{\alpha}_e}{\bar{\alpha}_i\beta_{in}{}^2 B^2}-\frac{1}{\beta_{ei}\beta_{in}B^2}\right]\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B}\times(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)+\mathbf{B}\times\bar{\mathbf{c}}_n \end{aligned} \tag{44}$$

Finally, the current $\bar{\mathbf{j}}$ is considered. From CHAPMAN and COWLING (1952, Ch. 18),

$$\bar{\mathbf{j}} = \rho_p \bar{\mathbf{c}}_w + \bar{\mathbf{j}}_{\text{cond}}$$

where $\bar{\mathbf{j}}_{\text{cond}} = \sum_s \bar{\rho}_s \bar{\mathbf{C}}_w$ and ρ_p is the plasma charge density. It is still assumed that the plasma is neutral at all times so that

$$\bar{\mathbf{j}} = \bar{\mathbf{j}}_{\text{cond}} = \bar{\rho}_i(\bar{\mathbf{c}}_i - \bar{\mathbf{c}}_e) \tag{45}$$

Substituting known expressions for $\bar{\mathbf{c}}_i$ and $\bar{\mathbf{c}}_e$ into equation (45), we obtain

$$\begin{aligned}\bar{\mathbf{j}}\bar{\rho}_i{}^{-1} = {} & \Gamma^{-1}\left(1-\frac{\beta_{en}}{\beta_{in}}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)_\parallel + \frac{\Gamma}{\beta_{en}B^2}\left(\frac{1}{\beta_{en}}-\frac{1}{\beta_{in}}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) \\ & - \frac{1}{B^2}\left(\frac{1}{\beta_{en}}-\frac{1}{\beta_{in}}\right)\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B}\times(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) \end{aligned} \tag{46}$$

or as an approximation valid throughout most of the ionosphere,

$$\bar{\mathbf{j}} \simeq \bar{\rho}_e\left[\Gamma^{-1}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\frac{\beta_{en}}{\beta_{in}}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)_\parallel + \frac{\Gamma}{\beta_{in}\beta_{en}B^2}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n) - \frac{1}{\beta_{in}B^2}\frac{\bar{\alpha}_w}{\bar{\alpha}_i}\mathbf{B}\times(\bar{\mathbf{c}}_w-\bar{\mathbf{c}}_n)\right] \tag{47}$$

3.1 *Conclusions*

Given $\bar{\mathbf{c}}_w$, $\bar{\mathbf{c}}_n$, β_{en}, β_{in}, β_{ie}, and the variation of $\bar{\rho}_e$ with height (see Section 5.1), we can in turn obtain the detailed runs of $\bar{\mathbf{c}}_i$, $\bar{\mathbf{c}}_e$, $\bar{\mathbf{E}}$ and $\bar{\mathbf{j}}$ as functions of height. The needed runs of $\bar{\mathbf{c}}_w$ and $\bar{\mathbf{c}}_n$ will, of course, depend on the particular atmospheric temperature profile. An appropriate model is an atmosphere with a uniform vertical gradient of temperature. The solution of the classical wave-equation including linear thermal effects has been given in a previous paper (WHITE, 1956) in terms of Bessel functions. For application to the present discussion, it would only be necessary to generalize† $T(r)$ to include the non-linear source function. The procedure would appear to be straightforward.

Formerly, the practice has been to assume either that the tidal velocity is independent of height or to assume some arbitrary height variation. Thus, for the shell model where a tidal velocity independent of height is used, the actual variation of $\bar{\mathbf{E}}$ and $\bar{\mathbf{j}}$ with height remain unknown. The present unified treatment of tidal–dynamo effects gives the detailed vertical and horizontal variations as well as the so-called distortion of the F-layer (see Section 5.1).

4. STABILITY CRITERION AND HARMONIC RESPONSE

It is interesting to examine the conditions under which the oscillations will or will not be stable. The present method for investigating stability consists in applying a small perturbation and examining the conditions for which the perturbations will grow or attenuate, the latter representing a stable condition, the former an unstable one. The equation (28) under investigation consists of a single dependent variable $\bar{\psi}_r$ and four independent variables, three in space, and one in time. Usually an analysis for stability is restricted to one- or two-dimensional space. It will be shown that for the particular wave-equation in question, a sufficient condition for stability is that the non-linearity be variables separable and specifically that

$$\bar{q}_{r(\nu,\eta)}(\theta,\phi,z,t) = q_{r;\nu,\eta}(\theta)\,e^{is\phi}\,q_{r;\nu,\eta}(z)\,\psi_\nu(t)\,\psi_\eta(t) \tag{48}$$

where $q_{r;\nu,\eta}(\theta)$ is a real function satisfying equation (16) of WILKES (1951), $q_{r;\nu,\eta}(z)$ is an arbitrary complex function, and the ϕ-dependence $q_{r;\nu,\eta}(\phi)$ equals $e^{is\phi}$, the ϕ-dependence for the classical tidal theory. The exponent $is\phi$ is purely imaginary. $\psi_\nu(t)$ and $\psi_\eta(t)$ represent $e^{i\sigma_\nu t}$ and $e^{i\sigma_\eta t}$, respectively, where the exponents are purely imaginary.

From Section 3, equations (40) and (41), it is seen that the non-linearities in time are of the third degree. It appears convenient to reduce the non-linearities from third to second degree [equation (42)] so that a representative term for $\bar{J}_r$ is given by

$$\bar{J}_r = \sum_{\nu=0}^{\infty} Q_{r,\nu}(\theta,\phi,z)\,\psi_\nu(t) + \sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} q_{r;\nu,\eta}(\theta,\phi,z)\,\psi_\nu(t)\,\psi_\eta(t) \tag{49}$$

consisting of the dominant linear source function, $Q_{r,\nu}(\theta,\phi,z)\,\psi_\nu(t)$ introduced by WILKES (1951) and a non-linear frictional term, $\bar{q}_{r;\nu,\eta}(\theta,\phi,z,t)$.

† See equation (22) of WHITE (1956).

VOL. 17 1959/60

Let us now substitute equation (49) and representative term of the linearized solution

$$\bar{\psi}_r = \sum_\nu \psi_{r,\nu}(\theta,\phi,z)\,\psi_\nu(t)$$

into equation (28), obtaining

$$\sum_{\nu=0}^{\infty} K_{r,\nu}{}^{\mathrm{I}}\ddot{\psi}_\nu(t) + \sum_{\nu=0}^{\infty} L_{r,\nu}{}^{\mathrm{I}}\dot{\psi}_\nu(t) + \frac{1}{2}\sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} L_{r(\nu,\eta)}{}^{\mathrm{II}}(\theta\phi z)\,\psi_\nu(t)\,\dot{\psi}_\eta(t) + \dot{\psi}_\nu(t)\,\psi_\eta(t)$$
$$+ \sum_{\nu=0}^{\infty} G_{r,\nu}{}^{\mathrm{I}}\psi_\nu(t) + \sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} G_{r(\nu,\eta)}{}^{\mathrm{II}}\psi_\nu(t)\,\psi_\eta(t)$$
$$= \sum_{\nu=0}^{\infty} J_{r,\nu}{}^{\mathrm{I}}\dot{\psi}_\nu(t) + \sum_{\nu=0}^{\infty} J_{r,\nu}{}^{\mathrm{II}}\psi_\nu(t) \tag{50}$$

where the coefficients K, L, G and J are time-independent operators on $\psi_{r,\nu}(\theta,\phi,z)$, $Q_{r,\nu}(\theta,\phi,z)$, or $q_{r;\nu,\eta}(\theta,\phi,z)$.

In order to examine the question of stability, let $\bar{\psi}_r = \sum_\nu \psi_{r,\nu}(\theta,\phi,z)\,\psi_\nu(t)$ be replaced by

$$\bar{\psi}_r = \sum_\nu \{\psi_{r,\nu}(\theta,\phi,z)\,[\psi_\nu(t) + \xi_\nu(t)]\} \tag{51}$$

where $\xi(t)$ represents the small arbitrary variation in the displacement. Equation (49) for $\bar{J}_r$ now becomes

$$\bar{J}_r = \sum_{\nu=0}^{\infty} Q_{r,\nu}(\theta,\phi,z)\,\psi_\nu(t) + \sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} q_{r;\nu,\eta}(\theta,\phi,z)\,[\psi_\nu(t)\,\psi_\eta(t) + \psi_\nu(t)\,\xi_\eta(t) + \psi_\eta(t)\,\xi_\nu(t)] \tag{52}$$

where the term $\xi_\eta(t)\,\xi_\nu(t)$ is assumed to be negligibly small. With the help of equations (51) and (52), the variational equation corresponding to equation (50) is formed, and then equation (50) itself is subtracted from the variational equation, giving

$$\sum_{\nu=0}^{\infty} K_{r,\nu}{}^{\mathrm{I}}\ddot{\xi}_\nu(t) + \sum_{\nu=0}^{\infty} L_{r,\nu}{}^{\mathrm{I}}\dot{\xi}_\nu(t) + \frac{1}{2}\sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} L_{r,(\nu,\eta)}{}^{\mathrm{II}}[\psi_\nu(t)\,\dot{\xi}_\eta(t) + \dot{\psi}_\nu(t)\,\xi_\eta(t) + \psi_\eta(t)\,\dot{\xi}_\nu(t) + \dot{\psi}_\eta(t)\,\xi_\nu(t)]$$
$$+ \sum_{\nu=0}^{\infty} G_{r,\nu}{}^{\mathrm{I}}\xi_\nu(t) + \sum_{\nu=0}^{\infty}\sum_{\eta=0}^{\infty} G_{r(\nu,\eta)}{}^{\mathrm{II}}[\psi_\nu(t)\,\xi_\eta(t) + \psi_\eta(t)\,\xi_\nu(t)] = 0 \tag{53}$$

One consequence of this approach is that the forcing functions J^{I} and J^{II} do not appear in equation (53) if they are purely external, so that the same perturbation equation holds for the force-free case (conduction and convection are therefore being neglected here). Representing $\psi_\nu(t)$ by

$$\psi_\nu(t) = \psi_{\nu c}\cos\nu\sigma_1 t + \psi_{\nu s}\sin\nu\sigma_1 t$$

where $\psi_{\nu c}$ and $\psi_{\nu s}$ are time-independent, equation (53) is seen to be a third-order linear differential equation in $\xi(t)$ with periodic coefficients. Floquet's theory (STOKER, 1950) is therefore applicable, and suggests a solution of the form

$$= \sum_{\eta=0}^{\infty} \exp(i\,\mathbf{a}_\eta t)(a_\eta \sin\eta\sigma_1 t + b_\eta\cos\eta\sigma_1 t) = \sum_{\eta=0}^{\infty}\xi_\eta(t) \tag{54}$$

Substituting the above expressions for $\psi_\nu(t)$ and $\xi_\eta(t)$ in equation (53), and treating terms in $\sigma_1{}^2$ and $\sigma_1{}^3$ as negligible and rearranging, yields the equation of condition which involves an infinite sum of terms in $\cos\nu\sigma_1 t\exp(i\mathbf{a}_\nu t)$ and $\sin\nu\sigma_1 t\exp(i\mathbf{a}_\nu t)$ as well as a doubly infinite sum of terms of the type $\cos\eta\sigma_1 t\cos\nu\sigma_1 t\exp(i\mathbf{a}_\eta t)$, etc., the sum of which equals zero.

A sufficient condition that the equation of condition be valid for all time is, in part,

$$\tfrac{1}{2}L_{r(\nu,c\eta)}{}^{\mathrm{II}}(\psi_{\nu c}\eta\sigma_1 a_\eta + \psi_{\nu c}\,i\mathbf{a}_\eta b_\eta + \psi_{\nu s}\nu\sigma_1 b_\eta) + G_{r(\nu,\eta)}{}^{\mathrm{II}}\psi_{\nu c}\,b_\eta = 0 \tag{55}$$

$$\tfrac{1}{2}L_{r(\nu,\eta)}{}^{\mathrm{II}}(-\psi_{\nu s}\eta\sigma_1 b_\eta + \psi_{\nu s}\,i\mathbf{a}_\eta a_\eta - \psi_{\nu c}\nu\sigma_1 a_\eta) + G_{r(\nu,\eta)}{}^{\mathrm{II}}\psi_{\nu s}\,a_\eta = 0 \tag{56}$$

from consideration of the coefficients to $\cos\eta\sigma_1 t\cos\nu\sigma_1 t$ and $\sin\eta\sigma_1 t\sin\nu\sigma_1 t$. Two similar conditions are obtainable from these same coefficients by simply exchanging the indices η and ν. From the coefficient for $\cos\eta\sigma_1 t\sin\nu\sigma_1 t$ and $\sin\eta\sigma_1 t\cos\nu\sigma_1 t$, the following sufficiency

conditions suggest themselves,

$$\tfrac{1}{2} L_{r(\nu,\eta)}{}^{\mathrm{II}} (\psi_{\nu s}\, \eta\sigma_1\, a_\eta + \psi_{\nu s}\, i\mathbf{a}_\eta\, b_\eta - \psi_{\nu c}\, \nu\sigma_1\, b_\eta) + G_{r(\nu,\eta)}{}^{\mathrm{II}} \psi_{\nu s}\, b_\eta = 0 \tag{57}$$

$$\tfrac{1}{2} L_{r(\nu,\eta)}{}^{\mathrm{II}} (-\psi_{\nu c}\, \eta\sigma_1\, b_\eta + \psi_{\nu c}\, i\mathbf{a}_\eta\, a_\eta + \psi_{\nu s}\, \nu\sigma_1\, a_\eta) + G_{r(\nu,\eta)}{}^{\mathrm{II}} \psi_{\nu c}\, a_\eta = 0 \tag{58}$$

where two similar conditions also obtain by simply reversing subscripts η and ν.

Multiplication of equation (55) by $\psi_{\nu c}$ and equation (57) by $\psi_{\nu s}$, then adding and dividing through by ${\psi_{\nu c}}^2 + {\psi_{\nu s}}^2$ leads to the equation of condition

$$\tfrac{1}{2} L_{r(\nu,\eta)}{}^{\mathrm{II}} (\eta\sigma_1\, a_\eta + i\mathbf{a}_\eta\, b_\eta) + G_{r(\nu,\eta)}{}^{\mathrm{II}} b_\eta = 0 \tag{59}$$

Similarly, by multiplying equations (56) and (58), respectively, by $\psi_{\nu s}$ and $\psi_{\nu c}$, adding, and dividing by ${\psi_{\nu s}}^2 + {\psi_{\nu c}}^2$, we have

$$\tfrac{1}{2} L_{r(\nu,\eta)} (-\eta\sigma_1\, b_\eta + i\mathbf{a}_\eta\, a_\eta) + G_{r(\nu,\eta)}\, a_\eta = 0 \tag{60}$$

which is compatible with equation (59) if we set

$$\frac{a_\eta}{b_\eta} = \pm i \tag{61}$$

There now remains only the consideration of the linear terms in the equation of condition. By setting the coefficients of $\cos \nu\sigma_1 t$ and $\sin \nu\sigma_1 t$ separately equal to zero and then multiplying one of the two equations by the imaginary i and adding, one obtains two final equations of condition, namely,

$$\left.\begin{aligned} \psi_{\nu c}\, a_\nu - \psi_{\nu s}\, b_\nu &= 0 \\ \psi_{\nu c}\, b_\nu + \psi_{\nu s}\, a_\nu &= 0 \end{aligned}\right\} \tag{62}$$

and

Equations (62) are, however, completely compatible with condition (61). It therefore appears that the two conditions (59) and (61) are sufficient to ensure that the equation of condition vanishes for all times; they may, in fact, be replaced by the single condition for $\mathbf{a}_\eta$ that

$$\mathbf{a}_\eta = \frac{2G_{r(\nu,\eta)}{}^{\mathrm{II}}}{L_{r(\nu,\eta)}{}^{\mathrm{II}}}\, i \mp \eta\sigma_1 \tag{63}$$

The non-linear oscillation will remain stable if ξ, given by equation (54), diminishes with time, or in other words,

$$\text{real part } (i\mathbf{a}_\eta) \leqslant 0 \tag{64}$$

Substitution of equation (63) with some slight reductions yields the sufficient condition for stability that

$$\text{real part } (G_{r(\nu,\eta)}{}^{\mathrm{II}} / L_{r(\nu,\eta)}{}^{\mathrm{II}}) \geqslant 0 \tag{65}$$

Applying the definitions for $G_{r(\nu,\eta)}{}^{\mathrm{II}}$ and $L_{r(\nu,\eta)}{}^{\mathrm{II}}$ and equation (16) of WILKES (1951), it is observed that the ratio (65) will be purely imaginary if equation (48) obtains, so that inequality (65) will be identically equal to zero. This is just the condition for neutral stability, and is interpreted as meaning that the oscillations will neither "blow up" nor die down to zero due to damping. The result seems reasonable in view of the assumption $\overline{\Pi}_r = \nabla\bar{\pi}_r$ used in the present treatment of stability; all collisional terms then take the role of a forcing function similar to $\bar{Q}$ or $\overline{\Omega}$ while actual damping effects are apparently neglected. Hence, the oscillations maintain at a constant amplitude, even for the force-free case [since forcing functions disappear in equation (53)]. We may therefore conclude that a sufficient condition for neutral stability obtains where the non-linear forcing function is variables separable and of the form given by equation (48).

VOL. 17 1959/6

4.1 *Harmonic response*

A formalized proof for the harmonic content follows from equation (50), which may be replaced by an infinite set of third-order time-dependent differential equations ($0 \leqslant \nu \leqslant \infty$). One then considers a representative equation, say the νth one. The method (STOKER, 1950, p. 83) consists in re-writing this νth equation as a solution for $\dddot{\psi}_\nu(t)$, thus

$$\ddot{\psi}(t) = -\mathscr{L}_{r,\nu}{}^{\mathrm{I}}\psi_\nu(t) - \mathscr{G}_{r,\nu}{}^{\mathrm{I}}\psi_\nu(t) + \mathscr{J}_{r\nu}{}^{\mathrm{I}}\dot{\psi}_\nu(t) + \mathscr{J}_{r\nu}{}^{\mathrm{II}}\psi_\nu(t)$$
$$-\frac{1}{2}\sum_{\eta=0}^{\infty}\mathscr{L}_{r(\nu,\eta)}{}^{\mathrm{II}}[\psi_\nu(t)\dot{\psi}_\eta(t)+\dot{\psi}_\nu(t)\psi_\eta(t)] - \sum_{\eta=0}^{\infty}\mathscr{G}_{r(\nu,\eta)}{}^{\mathrm{II}}\psi_\nu(t)\psi_\eta(t) \tag{66}$$

where $\mathscr{L} = L/K^{\mathrm{I}}$, $\mathscr{G} = G/K^{\mathrm{I}}$, and $\mathscr{J} = J/K^{\mathrm{I}}$ with corresponding sub- and superscripts. The first four terms on the right represent the linearized case, giving rise to no new harmonics. Now representing $\psi_\nu(t)$ by $\exp(\pm i\nu\sigma_1 t)$ (and likewise $\psi_\eta(t)$ by $\exp[i\eta\sigma_1 t]$), equation (66) may be easily integrated three times. The constants of integration are set equal to zero as we are interested in periodic solutions only. We therefore obtain as a higher approximation to $\psi(t)$,

$$\psi(t) \propto \sum_{\eta=0}^{\infty}\left\{[\tfrac{1}{2}\mathscr{L}_{r(\nu,\eta)}{}^{\mathrm{II}}(i\eta\sigma_1 \pm i\eta\sigma_1) + \mathscr{G}_{r(\nu,\eta)}{}^{\mathrm{II}}]\frac{\exp[i(\eta\pm\nu)\sigma_1 t]}{i(\eta\pm\nu)^3\sigma_1{}^3}\right\} \tag{67}$$

where now the combination "tones" $(\eta \pm \nu)$ appear. For $\eta = \nu$, one has self-coupling.

It thus appears that the non-linearity has the effect of making the oscillations harmonically richer.

5. Auxiliary Equations and Supporting Observational Evidence

Before proceeding to possible supporting observational evidence, it might be profitable to examine in brief the role of the daily variations of electron density; from examination of equations (40) and (41) it is seen that the non-linear terms $\bar{q}_r$ depend upon $\bar{\rho}_e$ so that for complete rigour one has to solve simultaneously the two non-linear wave-equations (25) and the equation of continuity (including transport term) for $\bar{\psi}_n$, $\bar{\psi}_w$, and $\bar{\rho}_e$. In what immediately follows, a solution to the general equation of electron continuity will be given by a method of successive approximation, where the zeroth order approximation is given by the usual steady state solution.

5.1 *Continuity equations for electrons*

The general equation for continuity of electrons is given by

$$\frac{\partial\bar{n}_e}{\partial t} = \bar{q} - a_\nu \bar{n}_e{}^\nu - \bar{\mathbf{c}}_e.\nabla\bar{n}_e - \bar{n}_e\bar{\chi}_e \tag{68}$$

We shall here exclude the possibility for both cases ($\nu = 1$ and 2) being important at the same station. In order to demonstrate the proposed method, we shall limit ourselves to the dominant components $\bar{c}_e \simeq \bar{c}_{e2}$ and $\bar{\chi}_e \simeq \bar{\chi}_{e2}$, where the subscript 2 indicates the semi-diurnal component.

The source function $\bar{q}_r$ or $\bar{Q}$ may be modelled according to

$$\bar{q}_r = \begin{cases} q_r(z)\cos\zeta & \text{days} \\ 0 & \text{nights} \end{cases} \tag{69}$$

where ζ = solar zenith angle.

This model may be profitably compared with the Chapman (MITRA, 1952) layer formation theory, where equilibrium conditions are assumed, or in effect

$$\bar{q} \propto \bar{n}_e{}^2 \propto \cos\zeta \tag{70}$$

valid for the maximum rate of ion production. The present model assumes equation (73) to hold for day conditions only, that is, for the solar control terms, it is supposed that a day-time Chapman layer is not too bad a representation. For the *F1*- and *F2*-regions, the Chapman layer will be "distorted" by the additional transport terms in equation (71) and by the day–night effect.

Function (69) can be analysed into spherical harmonics as follows:

$$\begin{aligned}\bar{q}_r = q_r(z)[\bar{q}_{r,0}(\theta,\phi) &+ 0{\cdot}5P_1^{\,1}(\theta)\sin(\omega t' + 270^\circ) + 0{\cdot}27P_2^{\,2}(\theta)\sin(2\omega t' + 90^\circ) \\ &+ 0{\cdot}052P_4^{\,2}(\theta)\sin(2\omega t' + 90^\circ) + 0{\cdot}069P_4^{\,4}(\theta)\sin(4\omega t' + 270^\circ) \\ &+ 0{\cdot}050P_6^{\,4}(\theta)\sin(4\omega t' + 270^\circ)] \quad \text{for} \quad \delta = 0^\circ, \end{aligned} \tag{71}$$

where the numerical coefficients and phases are given by KERTZ (1956b) (for a different application); t' is local time measured from midnight and δ is the solar declination. For solstice conditions ($\delta = 23^\circ$), $P_2^{\,1}(\theta)$, $P_4^{\,1}(\theta)$ and $P_4^{\,3}(\theta)$ terms also make slight contributions. The value of $\bar{q}_{r,0}$ is not presently available, but we can get some information on its value by noting that $\bar{q}_{r,0}$ cannot be negative, that is,

$$\bar{q}_{r,0} \geqslant 0{\cdot}5P_1^{\,1}(\theta) + 0{\cdot}27P_2^{\,2}(\theta) + 0{\cdot}052P_4^{\,2}(\theta) + \ldots$$

or

$$\bar{q}_{w,0} \cong \bar{q}_{n,0} \cong q_0 = q_{0,0} + q_{1,0}P_1^{\,1}(\theta) + q_{2,0}P_2^{\,2}(\theta) + \ldots \tag{72}$$

where the $q_{0,0}$, $q_{1,0}$, $q_{2,0}$-terms are independent of co-latitude [see equation (30)]. In what follows, all the time-dependent variables are expanded in Fourier series.

For the case $\nu = 1$, we have for the steady state [from equation (68)]

$$\bar{n}_{e,0} = q_0/a_1 \tag{73}$$

for the diurnal component

$$\bar{n}_{e1} = \frac{(a_1 - i\sigma_1)\,\bar{q}_1}{a_1^{\,2} + \sigma_1^{\,2}} \tag{74}$$

for the semi-diurnal components, from equations (68) and (73),

$$\bar{n}_{e,2} = \frac{a_1 - 2i\sigma_1}{a_1^{\,2} + 4\sigma_1^{\,2}}\left[\bar{q}_2 - \bar{\mathbf{c}}_{e,2}\cdot\left(\frac{1}{a_1}\nabla q_0 - \frac{q_0}{a_1^{\,2}}\nabla a_1\right) - \frac{q_0}{a_1}\bar{\chi}_{e,2}\right] \tag{75}$$

Again, from equation (68), one obtains for the case $\nu = 2$ and the steady state

$$\bar{n}_{e,0} = (q_0/a_2)^{\frac{1}{2}} \tag{76}$$

for the diurnal component, from equations (68) and (76),

$$\bar{n}_{e,1} = \bar{q}_1/[i\sigma_1 + 2\sqrt{(a_2 q_0)}] \tag{77}$$

and for the semi-diurnal component of $\bar{n}_e$, from equations (68), (76), and (77),

$$\begin{aligned}\bar{n}_{e\,2} = \frac{2\sqrt{(a_2 q_0)} - 2i\sigma_1}{4a_2 q_0 + 4\sigma_1^{\,2}}\Bigg\{\bar{q}_2 - \frac{[2\sqrt{(a_2 q_0)} - i\sigma_1]^2}{(4a_2 q_0 + \sigma_1^{\,2})^2}\,a_2\bar{q}_1^{\,2} \\ - \bar{\mathbf{c}}_{e,2}\cdot\nabla(q_0/a_2)^{\frac{1}{2}} - (q_0/a_2)^{\frac{1}{2}}\bar{\chi}_{e,2}\Bigg\}\end{aligned} \tag{78}$$

For both cases, $\nu = 1$ and 2, a model is needed for $a_\nu(z)$, $q_0(z)$, and $q_2(z)$; $q_1(z)$ is needed for the case $\nu = 2$.†

† This outline of a general solution to the continuity equation may be profitably compared, for example, with the method of E. V. APPLETON and A. J. LYONS (1954) *Report of Conference on the Physics of the Ionosphere* p. 20. Physical Society, London.

VOL. 17 1959/6

5.2. *Temperature barriers and boundary conditions*

WEEKES and WILKES have drawn a comparison between the "tidal" wave-equation:

$$\frac{d^2y}{dx^2}+\left[-\frac{1}{4}+\frac{1}{h}\left(\frac{dH}{dx}+\frac{\gamma-1}{\gamma}H\right)\right]y=0 \tag{79}$$

and the wave-equation for electromagnetic propagation

$$\frac{d^2E}{dx^2}+\mu^2E=0$$

with the net result that the coefficient of y in equation (82) is identified with the square of the index of refraction μ [see pp. 48–49 of WILKES (1949)]. WEEKES and WILKES are then able to conclude that for regions of the atmosphere where

$$\left[-\frac{1}{4}+\frac{1}{h}\left(\frac{dH}{dx}+\frac{\gamma-1}{\gamma}H\right)\right]\leqslant 0$$

that is, where there is a sufficient decrease in temperature, the atmosphere acts as a barrier to "tidal" energy flow.

Let us reconsider this identification of the coefficient of y in equation (79) with the index of refraction. The equation (28) for tidal motion is of third order in time; the second time derivation does not appear, probably due to the neglect of vertical acceleration (long wave approximation). It would therefore seem inadvisable to compare the wave-equation for electromagnetic waves or for sound waves with the so-called tidal wave-equation without further analysis.

Current observational evidence seems to substantiate the use of an atmospheric "top" with a constant positive temperature gradient. Physically speaking, the temperature rise in the F-region and above should act as a barrier to oscillations originating from below so that the earth's atmosphere will behave as a cavity resonator with a moderate Q. For the modes of vibration for which this holds, the high ionization content beyond the barrier will be unimportant and present treatment with the assumption of a slightly ionized gas throughout is justified. Conversely, oscillations originating from above 400 km can easily penetrate, as such oscillations will "see" a temperature decrease, at least down to the 80–100 km level. As a boundary condition for the atmospheric top, I would therefore suggest, in contrast with pp. 32–33 of WILKES (1949), that at sufficiently high levels, the energy flow be in a downward direction.

In a previous paper by the writer (WHITE, 1956) on "tidal" oscillations, the concept of an outgoing and incoming wave was extended to include outgoing and incoming Hankel functions. A further generalization of this concept is useful for solution of the inhomogeneous wave-equation (5) from that work, repeated here

$$\frac{d^2y}{dx^2}+\mu^2y=T(x) \tag{80}$$

where the homogeneous solution was given by

$$y=(P+iQ)(y_1+iy_2)+(R+iS)(y_1-iy_2)$$

Application of the method of variation of parameters leads to the following solution to (80):

$$y = \left\{(P+iQ) + \int \frac{(y_1 - iy_2)\,T(x)}{2i[y_1(dy_2/dx) - y_2(dy_1/dx)]}\,dx\right\}[y_1 + iy_2]$$
$$+ \left\{(R+iS) - \int \frac{(y_1 + iy_2)\,T(x)}{2i[y_1(dy_2/dx) - y_2(dy_1/dx)]}\,dx\right\}[y_1 - iy_2] \qquad (81)$$

where we interpret $[y_1 + iy_2]$ as a generalized outgoing wave and $[y_1 - iy_2]$ as a generalized incoming wave. Equation (81) or equations (7) and (8) of WHITE (1956) should be applicable to the present discussion provided $T(x)$ is defined so as to include the non-linear source function (compare with Section 3.1).

5.3. *Comparison of theory and observed tidal pressure variations*

As shown in equation (8) of the first section, ordinary dynamical viscosity can produce a non-linear viscous heating term of the second degree, so that lower non-ionized atmospheric regions may give rise to non-linear vibrations, apart from the D/Dt-terms in the equations of motion. As check, the New York University's harmonic analysis (KERTZ, 1956a) of the 6 hr pressure variation $W_4{}^4$ is considered. The 6 hr pressure wave consists of many components

$$W_4{}^4 = \sum_{\eta=4}^{\infty} T_{4,\eta}{}^4$$

where, following KERTZ,

$$T_{\sigma,\eta}{}^s = \text{const}\, P_\eta{}^s(\theta) \sin(\sigma t^* + s\phi + \epsilon)$$

where t^* = Greenwich time. The most important term is $T_{4,4}{}^4$ with $\epsilon = 225°$. The corresponding h-value is 15·3 km, which is apparently far removed from resonance conditions ($h \simeq 7{\cdot}9$ km). It is difficult to explain this term on the basis of the so-called resonance theory [recall interpretation for equation (24)] unless there is a non-linear coupling of the important $T_{2,2}{}^2$ term with itself, that is, a self-coupling of the 12 hr oscillation to produce the important 6 hr component $T_{4,4}{}^4$. Let us recall that

$$T_{2,2}{}^2 = \text{const}\, P_2{}^2(\theta) \sin(2t' + 154°) \qquad (82)$$

If

$$T_{4,4}{}^4 = \text{const}\,[T_{2,2}{}^2]^2, \text{ then from equation (82)}$$
$$T_{4,4}{}^4 \propto \text{const}\, P_4{}^4(\theta) \sin(4t' + 218°) \qquad (83)$$

in close agreement with the experimental value of KERTZ and in support of conclusions stated above.†

It is tempting to try to explain the puzzling $W_2{}^3$ semi-diurnal migrating wave with three maxima and three minima around the earth's circumference, obtained by KERTZ (1956c), as a coupling between the progressive wave $T_{2,2}{}^2$ and a heat source fixed with respect to the earth's surface (arising from a land–sea distribution). For reference,

$$W_2{}^3 \propto P_3{}^3(\theta) \sin(2t^* + 3\phi + 26°)$$

and the corresponding $h_3 = 4{\cdot}60$ km.

† Skin friction at the earth's surface would also contribute to the non-linearity, but presumably to a lesser extent.

VOL. 17 1959/6

There is one final possibility for coupling that should be mentioned in closing, namely, that even in the absence of a semi-diurnal "solar" source function, non-linear effects could similarly cause the diurnal oscillation to excite a (2, 2) mode.

5.4. *Comparison of theory and Cavendish Tables* (THOMAS and ROBBINS, 1955) *of electron densities*

Direct comparison of theory and ionospheric observation is less satisfactory. Fortunately, the Cavendish Laboratory (THOMAS and ROBBINS, 1955) has recently published tables of the time variations of electron densities at given true heights for various months and stations, offering the possibility for indirect comparison. A

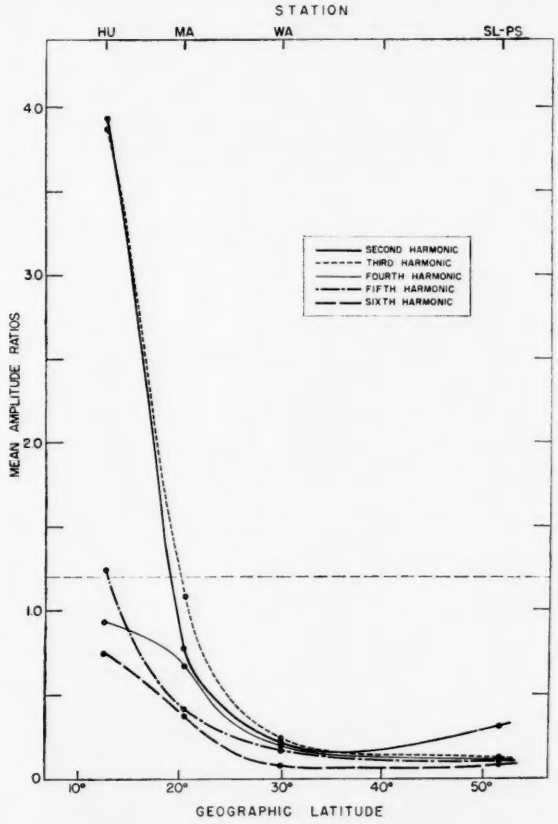


Fig. 1. Harmonic content of the daily variation in electron density (using Cavendish tables) at a given height (280–300 km) at Huancayo (HU), Maui (MA), Watheroo (WA), Slough (SL), and Port Stanley (PS), expressed as ratios to the first harmonic. Weighted mean values for the first harmonic of the electron density in units of 10^5 cm^{-3} are 1·26185 (HU), 2·69474 (MA), 2·95497 (WA), 3·70578 (SL), and 3·25484 (PS).

harmonic analysis of the electron density variation was performed on the National Bureau of Standards IBM 650 electronic computer out to the twelfth harmonic for a fixed height (280–300 km), for the four cardinal months, and for the five stations:

Maui, Port Stanley, Slough, Watheroo and Huancayo. Following this analysis, weighted mean amplitude ratios of the higher harmonics (2–6) to the first were plotted against latitude (see Fig. 1); when certain of the equinoctial (or solstitial) months were missing, others were given a correspondingly greater weight. From Section 5.1, on the continuity equation, it is apparent that in equatorial regions the large ratio of second and third harmonics to the fundamental (factor of 3 or 4) cannot be accounted for by the source function, which according to equation (71) would give a ratio of second to first harmonic of $\frac{1}{2}$; the harmonic richness can, however, be explained through the non-linearities of the transport terms. For ionospheric regions, the non-linearities are of the third degree in time [see equation (9)] as contrasted with the second degree non-linearities of the single component neutral gas atmosphere (applicable to Section 5.3) so that the third, as well as the second harmonic is here expected to be important. The representation of $\bar{Q}$, however, by a rapidly convergent series (71) discourages the appearance of still higher harmonics. By the same reasoning non-linearities produced by recombination appear inadequate for explaining the observations, and further, electron attachment seems presently favoured over recombination. We may therefore recognize in these ionospheric observations, possible I^2R non-linear effects.

At intermediate latitudes (see Fig. 1), the electron density variations are adequately explained through $\bar{q}$. At higher latitudes, the harmonic content increases again, at least for the ratio of second to first harmonic and it would be interesting to have data for even higher latitudes.

It is encouraging to note that when the data are halved and the earlier years are compared with the later years, Fig. 1 is recovered.

The present material was first announced in a letter to the editor (WHITE, 1957)† and copies of a manuscript dealing with the present subject matter were distributed at the URSI, XII General Assembly, Boulder, Colorado, 22 August–5 September, 1957.

Acknowledgements—It is a pleasure to acknowledge helpful conversations with Mr. V. AGY, Dr. H. H. HOWE, and Dr. E. L. CROW of the National Bureau of Standards, and to thank Mrs. A. HESSING for the computations on the IBM 650, and F. J. BURMONT for hand computations.

APPENDIX 1

Approximations leading to a generalized Kirchhoff's voltage law

The examination of the relative importance of electrical and mechanical terms for $r = e, i$ in the basic equation of motion (5) proceeds as follows: Let

$$\gamma_r = \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{\nabla p_r} \right| \simeq \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{p_r/H} \right| \gg \left| \frac{\bar{\mathbf{c}}_r \times \mathbf{B} \bar{\rho}_r H}{\bar{p}_r} \right| = \left| \frac{e}{m_r} \bar{\mathbf{c}}_r \times \mathbf{B} \frac{\bar{\alpha}_r H}{\bar{p}_r} \right| \tag{A1}$$

or

$$\gamma_r \gg \left| \left(\frac{e}{m_r} \right) \bar{\mathbf{c}}_r \times \mathbf{B}/g \right| \geqslant \left| \left(\frac{e}{m_r} \right) \bar{\mathbf{c}}_w \times \mathbf{B}/g \right|$$

† There are a few serious typographical errors in this letter to the editor.

VOL. 17 1959/6

or neglecting angular dependence and setting $|\bar{\mathbf{c}}_w| \simeq 5 \times 10^3$ cm/sec and $|\mathbf{B}| \simeq \frac{1}{2}$ gauss

$$\gamma_e \gg 4{\cdot}4 \times 10^7 \text{ and } \gamma_i \gg 830$$

Similarly, let

$$\delta_r = \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{\alpha_r g} \right| \gg \left| \frac{\bar{\rho}_r}{\bar{\alpha}_r} \frac{\bar{\mathbf{c}}_r \times \mathbf{B}}{g} \right| = \left| \left(\frac{e}{m_r} \right) \frac{\bar{\mathbf{c}}_r \times \mathbf{B}}{g} \right|$$

Comparing with γ_r, we again obtain

$$\delta_e \gg 4{\cdot}4 \times 10^7 \quad \text{and} \quad \delta_i \gg 830$$

and identical results for

$$\epsilon_r = \left| \frac{\bar{\rho}_r \bar{\mathbf{c}}_r \times \mathbf{B}}{\bar{\alpha}_r \nabla \Omega} \right| \gg \left| \frac{\bar{\rho}_r}{\bar{\alpha}_r} \frac{\bar{\mathbf{c}}_r \times \mathbf{B}}{g} \right|$$

where $g \gg |\Delta\Omega|$.

Finally, terms to the left of the equality sign in basic equation (5) are immediately seen to be negligible due to (1) the angular frequencies σ and ω which are very small compared to one, and (2) the factor of specific mass $(e_r/m_r)^{-1}$.

Appendix 2

Computation of β_{en}, β_{in} and β_{ei} values

In Section 3, certain approximations are made which depend on the values of β_{en}, β_{in} and β_{ei}, defined in equations (32) and (43) of the text. It will be shown that the β-values reduce to

$$\beta_{en} \cong -e/m_e \nu_{en}, \quad \beta_{in} \cong e/m_i \nu_{in}, \quad \text{and} \quad \beta_{ei} = -\beta_{ie} = -e/m_e \nu_{ei} \tag{A2}$$

by employing the following approximations to $\bar{\theta}_{rs}$: following Cowling (1945), Hirono (1955) and Nicolet (1952):

$$\bar{\theta}_{rs} = \frac{\bar{n}_r \bar{n}_s}{\bar{n}_r + \bar{n}_s} \frac{kT_0}{D_{rs}} = \frac{\bar{\alpha}_r \bar{\alpha}_s}{\bar{\alpha}_r + \bar{\alpha}_s} \cdot \frac{1}{\tau_{rs}} \tag{A3}$$

where τ_{rs} is the collision interval. Since the region of the atmosphere under consideration is composed predominantly of neutral particles,

$$\bar{\theta}_{rn} \simeq n_r m_r \nu_{rn} \tag{A4}$$

where

$$\nu_{rn} = kT_0/m_r D_{rn} \quad \ldots (r = i, e) \tag{A5}$$

where ν_{in} and ν_{en} are a close approximation to the collisional frequencies of positive ions and electrons with neutrals; finally,

$$\bar{\theta}_{ie} \simeq n_e m_e \nu_{ie}$$

where

$$\nu_{ie} = kT_0/2m_e D_{ie} \tag{A6}$$

Formulae and methods for obtaining the ν-values in Table 1 depend on the work of Nicolet (1952). The atmospheric model is based on one given by Chapman.† Results are summarized in Table 1.

† Table 1 is taken in part from a class problem for a course offered by S. Chapman, Sept. 1955–Jan. 1956, sponsored by the National Bureau of Standards, Boulder Laboratories and the University of Colorado.

Table 1. Collision frequencies, β-values and other characteristics for a cool, c, and hot, h, ionospheric model above White Sands, N.M.

Height (km)	100	275	Height (km)	100	275
Mol. wt.	29	20	ν_{in} [h]	8590	1·7
n_e	10^5	$2{\cdot}75 \times 10^5$	ν_{in} [c]	8840	0·116
$T°$K [h]	280	1330	β_{en} [day]*	-110	$-3{\cdot}078 \times 10^5$
$T°$K [c]	250	775	β_{en} [night]	-113	$-5{\cdot}88 \times 10^6$
n_n [h]	$1{\cdot}78 \times 10^{13}$	$2{\cdot}92 \times 10^9$	β_{in} [d]	$3{\cdot}88 \times 10^{-2}$	195·8
n_n [c]	$1{\cdot}83 \times 10^{13}$	$2{\cdot}00 \times 10^8$	β_{in} [n]	$3{\cdot}77 \times 10^{-2}$	2870
ν_{ei} [h]	935	286	β_{ei} [d]	$-1{\cdot}89 \times 10^4$	$-6{\cdot}189 \times 10^4$
ν_{ei} [c]	1090	605	β_{ei} [n]	$-4{\cdot}06 \times 10^5$†	$-2{\cdot}93 \times 10^5$
ν_{en} [h]	$1{\cdot}61 \times 10^5$	57·5			
ν_{en} [c]	$1{\cdot}56 \times 10^5$	3·01			

* "Day" and "night" models are in part taken, respectively, from "hot" and "cool" models.

† According to JAMES WATTS, of the National Bureau of Standards, experiments at low frequencies show that the electron density decreases by a factor of approximately 25 in the E-region from day to night.

VOL.
17
1959/6

REFERENCES

BAKER W. G. and MARTYN D. F. 1953 *Phil. Trans.* A **246**, 281.

CHAPMAN S. 1956 *Nuevo Cim.* Suppl. **4**, 1385.

CHAPMAN S. and BARTELS J. 1940 *Geomagnetism* Vol. II, p. 768. Clarendon Press, Oxford.

CHAPMAN S. and COWLING T. G. 1952 *The Mathematical Theory of Non-uniform Gases.* Cambridge University Press.

COWLING T. G. 1945 *Proc. Roy. Soc.* A **183**, 453.

1956 *Mon. Not. R. Astr. Soc.* **116**, 114.

FARADAY M. 1832 *Phil. Trans. Roy. Soc.* **122**, 177.

HINES E. O. 1953 *J. Geophys. Res.* **58**, 95.

HIRONO M. 1955 *Rep. Ionos. Res. Japan* **9**, No. 2, 95.

JOHNSON M. H. 1951 *Phys. Rev.* **84**, 566.

KERTZ W. 1956a *The Seasonal Variations of the Six-hourly Planetary Pressure and Temperature Waves.* Scientific Report No. 5, Research Division, College of Engineering, New York University, Project No. 299.

1956b *Nachr. Akad. Wiss. Göttingen. Math-Phys-Chem. Abt.* **6**, 145.

1956c *Components of the Semi-diurnal Pressure Oscillation.* Scientific Report No. 4, Research Division, College of Engineering, New York University, Project No. 299.

McNish A. G.	1939	*Terrestrial Magnetism and Electricity* p. 345. McGraw-Hill, New York.
Miller J. E.	1951	*Compendium of Meteorology* (Edited by Malone T. F.) p. 483. American Meteorological Society.
Mitra S. K.	1952	*The Upper Atmosphere* p. 280. Asiatic Society.
Nicolet M.	1952	Scientific Report No. 32. Ionospheric Research Laboratory, Pennsylvania State University.
Sen H. K. and White M. L.	1955	*J. Geophys. Res.* **60**, 483.
Stewart B.	1882	*Encyclopaedia Britannica* (9th Ed.) p. 36.
Stoker J. J.	1950	*Nonlinear Vibrations in Mechanical and Electrical Systems* p. 193. Interscience, New York.
Thomas J. O. and Robbins A. R.	1955	*Tables of* F2*-Layer Electron Density on International Quiet Days.* Ser. A, Nos. 1 and 2. Cavendish Laboratory, Cambridge.
White M. L.	1956	*J. Geophys. Res.* **61**, 489.
	1957	*J. Geophys. Res.* **62**, 329.
Wilkes M. V.	1949	*Oscillations of the Earth's Atmosphere*, Ch. 2. Cambridge University Press.
	1951	*Proc. Roy. Soc.* A **207**, 358.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 246 to 253. Pergamon Press Ltd. Printed in Great Britain

An attempt to measure the collision frequency of electrons in the *F*-region of the ionosphere

D. M. SCHLAPP
Cavendish Laboratory, Cambridge*

(*Received* 21 *May* 1959)

Abstract—An attempt was made to measure the collision frequency (ν) of electrons in the *F*-region of the ionosphere by APPLETON's method of observing changes in the logarithmic reflection coefficient $\log\rho$ associated with changes ($\Delta P'$) in the group path, and using the expression

$$\Delta \log \rho = -\frac{\nu}{2c}(\Delta P' - \Delta P)$$

It is shown that the experimental results were very variable and that the average of a large number behaves in a way which is not in accord with theory. It is concluded that, in the *F*-region, ν is less than about 5×10^3 sec^{-1} but that no reliable estimates of its magnitude can be made by this method.

1. INTRODUCTION

THERE have been numerous reports of attempts to measure the collision frequency (ν) in the *F*-region (e.g. ECKERSLEY, 1935; FARMER and RATCLIFFE, 1935; WHITE and BROWN, 1936; ARGENCE and RAWER, 1956). Most of these have used APPLETON's method (1935) of comparing changes in the logarithmic reflection coefficient $\log\rho$ with the corresponding changes in group path (P') and phase path (P) and using the expression

$$\Delta \log \rho = -\frac{\nu}{2c}(\Delta P' - \Delta P) \qquad (1)$$

It is difficult to make experiments of this kind because the downcoming wave is always found to be fading. In the work to be described here, a rapid frequency change apparatus due to BRIGGS (1951) was used, which exhibited the amplitude of the reflected pulse and the corresponding echo delay on Cartesian co-ordinate axes on a cathode ray oscilloscope. The necessary frequency change could be made in 1–2 sec, during which time no appreciable fading had taken place.

2. EXPERIMENTAL MEASUREMENTS

A series of records made at intervals of 1 min is reproduced in Fig. 1, in the form of graphs of $\log\rho$ plotted against P'.† The variation in the shape of the curves is associated with frequency-selective fading, presumably caused by "focusing" by

* Now at University College of North Staffordshire.

† In obtaining values of ρ the amplitude of the reflected pulse was corrected for the effect of distance attenuation by assuming that the amplitude falls off inversely as $\int ds/\mu$ where ds is an element of path and μ is the refractive index.

curvature of the surfaces of constant electron density in the F-region (ARGENCE and RAWER, 1956). This shows that it is futile to attempt to obtain a value of ν from an individual plot of $\log \rho$ against P'.

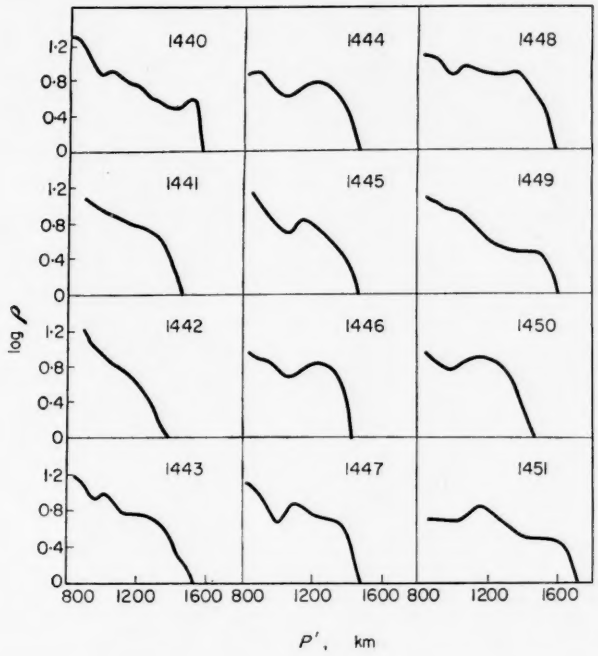


Fig. 1. Records showing the variation of the logarithmic reflection coefficient $\log \rho$ with group path P' for waves reflected from the $F2$-layer. The zero on the scale of $\log \rho$ is arbitrary.

In order to try to eliminate the effects of fading, it seems reasonable to take the average of a large number of such curves. One method of doing this is to take the average of several readings of amplitude at each of a series of values of group path. In an extensive series of measurements of this type carried out by the writer during 1955 and 1956 the readings were analysed in this way, but, even then, the values of ν showed various anomalies, which suggested that they were not reliable.

One such anomaly concerns the way in which ν varies with height (h) and electron density (N). These variations are shown in Figs. 2(a) and 2(b). It is seen that ν does not alter significantly over the range of h studied, and varies by a factor of at most 2 over the range of N from 10^5 cm^{-3} to 10^6 cm^{-3}.

The theoretical relation for collision frequency is

$$\nu = k_a n_a T^{\frac{1}{2}} + k_i n_i T^{-1} \qquad (2)$$

(COWLING, 1945) where n_a, n_i are the number densities of neutral particles and ions, respectively, T is the absolute temperature and k_a, k_i are constants involving,

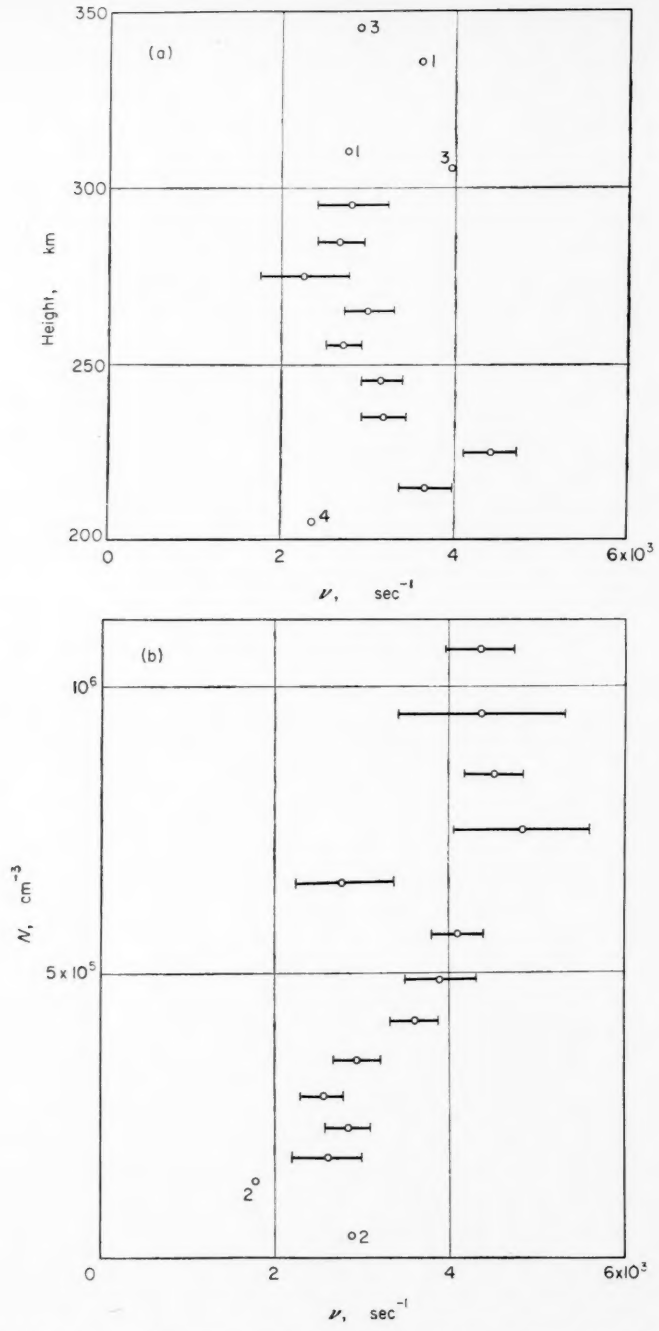


Fig. 2 (a) Variation of the measured values of ν in the *F2*-region with height. The error bars represent standard errors of the means. The number of observations which a point represents is given instead, where this was less than five. (b) Variation of the measured values of ν in the *F2*-region with peak electron density. The error bars represent standard errors of the means. The number of observations which a point represents is given instead, where this was less than five.

respectively, the collision cross-sections of neutral particles and of ions with electrons. Hence ν would be expected to vary either with neutral particle density and hence with height, or with ion density and hence with electron density, or with both height and electron density. Since the experimental values do not vary significantly with height over the wide range from 200 km to 350 km, it might appear that the first term on the right-hand side of equation (2) was negligible compared with the second. In this case ν ought to be proportional to electron density. It is possible that the observed variation with electron density could be explained by invoking a variation of temperature, but the results were sufficiently hard to explain for one to wonder whether the values were reliable.

When results for the *F1*- and *F2*-layers are compared, another anomaly is found. The mean of the values of ν measured for the *F1*-layer was significantly smaller than for the *F2*-layer for the same range of critical frequency. If collisions were predominantly with neutral particles, one would expect the value of ν in the *F1*-region to be greater than that in the *F2*-region, since the density of neutral particles decreases with height. If collisions were predominantly with ions, values of ν for the two layers for the same critical frequency should be similar. Hence the experimental result is again surprising.

These considerations make it seem likely that this type of measurement yields unreliable values of ν. It is, however, possible to deduce an upper limit for ν. The greatest average slope of the graphs of $\log \rho$ against P' corresponds to a value of ν of 5×10^3 sec^{-1}. Hence the collision frequency cannot be appreciably greater than this value, though it might be considerably less, if there is another source of attenuation besides deviative absorption.

OL.
17
9/60

3. The Shape of the Curve Relating $\log \rho$ to P'

Although equation (1) has usually been employed in measurements of ν it is, in fact, correct only if ν is independent of height. Near the penetration frequency of the F-layer the height of reflection changes by a considerable amount and it is desirable to investigate how far equation (1) might be expected to be true in the two cases where (a) collisions were with neutral particles or (b) with ions. In the interpretation of the results it is usual to neglect ΔP compared with $\Delta P'$ in equation (1). It is also desirable to investigate how far this neglect is justified. All these points are discussed in the Appendix, where it is shown that, for the range of frequencies used in this work, the curves relating $\log \rho$ to P' should be quite closely straight lines, whose slopes should give the value of ν at the peak of the layer accurate to about ± 10 per cent, even if ΔP is neglected in the calculation, and independent of whether collisions are with neutral particles or ions.

An examination of the individual experimental curves relating $\log \rho$ to P' shows that they tend to have a greater slope at larger values of group path, where the amplitude becomes small. This tendency is blurred out when the results are averaged over any useful time because the group path at which the slope becomes greater usually fluctuates considerably during that time. This fluctuation can be seen in Fig. 1.

To see whether the experimental curve departs systematically from a straight line when the amplitude becomes small, the averaging was done differently, as follows. The point where $\log \rho$ fell to some convenient arbitrary value was taken as a new origin on the group path axis and all curves were superposed so that they had this point in common. The observed values of $\log \rho$ were then averaged. The result is shown in Fig. 3, which indicates a systematic departure from a straight

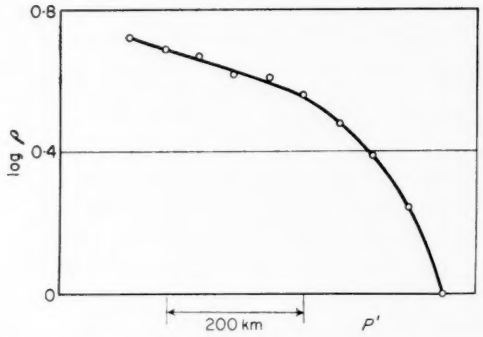


Fig. 3. Average curve of $\log \rho$ against P' for the *F2*-layer, derived by superposition as described in the text. The zero on the scale of $\log \rho$ is arbitrary.

line. This feature is always found when a series of curves of $\log \rho$ against P' for the *F2*-layer is analysed in this way. It is also found in the results of experiments made on the *F1*-layer.

4. Discussion

The systematic departure from a straight line in the curve relating $\log \rho$ to P' mentioned above is not in accordance with simple theory. Hence it must be concluded that measurements of this type for the *F*-region do not yield the correct value of ν, and it is not surprising that the anomalies mentioned in Section 2 were found.

When the type of analysis described above is applied to measurements on waves reflected from the *E*-region, it is found that the mean curve is straight over nearly the whole of its length, with a slight tendency to curvature only for the smallest amplitudes. Hence meaningful values of ν can be deduced by this method for *E*-region.

It appears likely that besides absorption due to collisions, there is another agency which causes the attenuation of highly retarded echoes reflected from the ionosphere, and that this attenuation is appreciable in the case of waves reflected from the *F*-region. Possible causes of this attenuation might be:

(a) Distance attenuation.

(b) A systematic variation in polarization with echo delay.

(c) Dispersion.

VOL.
17
1959/6

(d) Partial transmission of waves through a layer.

(e) Scattering caused by inhomogeneities in electron density.

All these possibilities have been carefully investigated and it is considered that none of them can satisfactorily explain the anomalous behaviour of the curve relating $\log \rho$ to P'.

5. Conclusions

The curves relating $\log \rho$ to P' show a systematic departure from the straight line expected on simple theory. Hence the values of ν obtained by measuring the slopes of these curves are unreliable.

It is concluded that, in the F-region, ν is less than about 5×10^3 sec^{-1} but that no more accurate estimate of its value can be made by this method.

Acknowledgements—The writer would like to express his gratitude to Mr. J. A. Ratcliffe and Dr. B. H. Briggs for the advice given in many helpful discussions. He is indebted to the Department of Scientific and Industrial Research for the award of a maintenance grant.

References

Appleton E. V.	1935	*Nature, Lond.* **135**, 618.
Argence E. and Rawer K.	1956	*Nuovo Cim.* **4**, Suppl. 4, 1511.
Briggs B. H.	1951	*Proc. Phys. Soc.* B **64**, 255.
Cowling T. G.	1945	*Proc. Roy. Soc.* A **183**, 453.
Eckersley T. L.	1935	*Nature, Lond.* **135**, 435.
Farmer F. T. and Ratcliffe J. A.	1935	*Proc. Roy. Soc.* A **151**, 370.
White F. W. G. and Brown L. W.	1936	*Proc. Roy. Soc.* A **153**, 639.

Appendix

This Appendix describes an investigation carried out by J. A. Ratcliffe.

In the work described above, the collision frequency (ν) was determined from the slope of the curve relating $\log \rho$ to P'. Two major approximations are involved in this, as follows:

(a) The change in phase path (ΔP) in equation (1) is neglected.

(b) Equation (1) is itself derived on the assumption that ν is independent of height.

In this Appendix, some results are given which indicate that the error incurred in making these approximations is likely to be small.

The model used in obtaining these results was a parabolic variation of electron density with height. The following cases were considered:

(a) ν independent of height.

(b) An exponential variation of ν with height as given by

$$\nu = \nu_0 \exp\left(-\frac{h}{H}\right)$$

This would be expected if collisions were with neutral particles.

(c) A variation of ν directly proportional to electron density as would be expected if collisions were with ions.

Then it can be shown that for case (a)

$$-\log \rho = \frac{\nu}{c}\frac{1}{f}A(f)$$

for case (b)

$$-\log \rho = \frac{\nu_p}{c}\frac{1}{f}B(f)$$

and for case (c)

$$-\log \rho = \frac{\nu_p}{c}\frac{1}{f}C(f)$$

where

$$A(f) = \int_1^a \frac{1-x^2}{\sqrt{(x^2-a^2)}}\,dx$$

$$B(f) = \int_1^a \frac{(1-x^2)\,e^{bx}}{\sqrt{(x^2-a^2)}}\,dx$$

$$C(f) = \int_1^a \frac{(1-x^2)^2}{\sqrt{(x^2-a^2)}}\,dx$$

and $f = \dfrac{\omega}{\omega_p}$;

ω is the angular frequency of the wave;

ω_p is the angular penetration frequency;

ν_p is the value of ν at the peak of the layer;

x is the height measured downwards from the centre of the layer in units of the semi-thickness of the layer;

$a = 1-f^2$ (so that $x = a$ is the height where the refractive index $\mu(f) = 0$);

b is the ratio of the semi-thickness of the layer to the scale height of ν.

For a parabolic approximation to a Chapman layer, $b = 2$.

The functions

$$\frac{1}{f}A(f), \quad \frac{1}{f}B(f) \quad \text{and} \quad \frac{1}{f}C(f)$$

are plotted against P' in Fig. 4, as curves (a), (b) and (c), respectively. For case (b) two curves are shown, corresponding to $b = 1$ and $b = 2$. In the notation of this Appendix,

$$P' = 2f\int_1^a \frac{dx}{\sqrt{(x^2-a^2)}}$$

Curve (a) shows the effect of plotting $\log \rho$ against P' instead of $(P'-P)$. Curves (b) and (c) show in addition the effect of a variation of ν with height.

In the experiments described above, the range of frequency covered corresponded at most to $f = 0{\cdot}9$–$1{\cdot}0$. On the normalized scale used in Fig. 4, this range corresponds

VOL.
17
1959/6

to $P' > 2{\cdot}6$. In this range, it is seen that all the curves are straight to a first approximation. This is the result quoted in paragraph (3). The values of the slopes

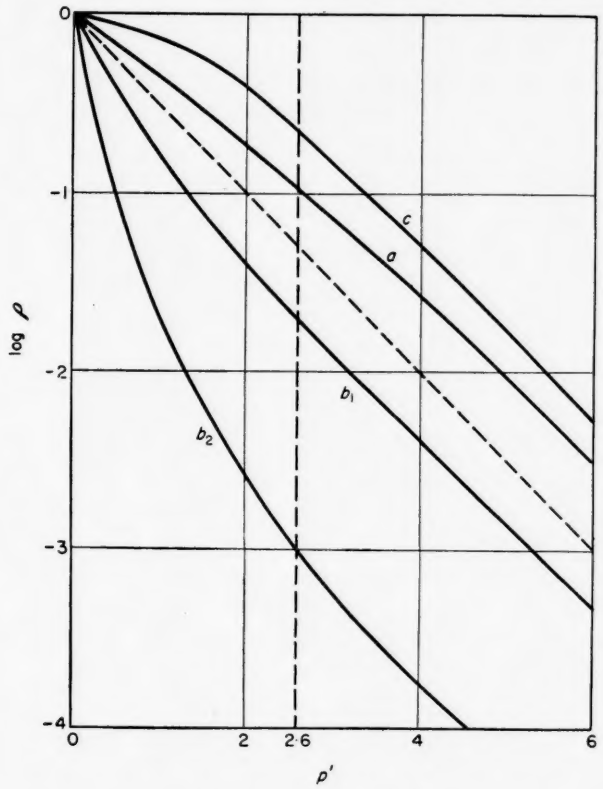


Fig. 4. Theoretical variation of $\log \rho$ with P', assuming:
(a) ν independent of height;
(b) an exponential variation of ν with height;
(c) a variation of ν directly proportional to electron density.

vary by about ± 10 per cent, which means that these approximations lead to errors of about this magnitude in the values of ν.

Papers to be published in future issues

R. L. DOWDEN: Ionospheric thermal radiation at radio frequencies in the auroral zone

G. R. A. ELLIS: Gyro splitting of ionospheric echoes

E. D. R. SHEARMAN and J. HARWOOD: Sporadic-*E* as observed by back-scatter techniques in United Kingdom

V. MARASIGAN: Spread-*F* in Baguio through half a solar cycle

T. OBAYASHI and Y. HAKURA: Enhanced ionization in the polar ionosphere associated with geomagnetic storms

R. A. DUNCAN: The equatorial *F*-region of the ionosphere

J. E. C. GLIDDON and P. C. KENDALL: Theoretical world curves of maximum ion density in a quiet *F*-region

J. A. GLEDHILL and A. D. M. WALKER: The "Valley effect" in the interpretation of ionospheric eclipse records

R. F. STEBBINGS, G. V. GROVES, R. L. F. BOYD and P. J. BOWEN: Low-altitude testing of the rocket-grenade experiment for upper atmosphere temperatures and winds

J. VELDKAMP and D. VAN SABBEN: On the current system of solar-flare effects

P. DAS GUPTA and K. K. VIJ: Statistical analysis of fading of a single down-coming wave

D. R. BATES and M. NICOLET: Ion–atom exchange

P. BANDYOPADHYAY: Electron distribution in the *F*-layer of the ionosphere over Haringhata (Calcutta) on quiet and disturbed days

J. TAUBENHEIM: Incoherency of pulse echoes and the measurement of ionospheric absorption

J. A. FEJER: Hydromagnetic wave propagation in the ionosphere

J. S. GREENHOW and J. E. HALL: Diurnal variations of density and scale height in the upper atmosphere

D. LEPECHINSKY and C. TAIEB: La détermination des hauteurs réelles de l'ionisation dans les très hautes latitudes

S. R. KHASTGIR and R. N. SINGH: The size of the moving irregularities on the *F*-region and the spread angle of the radio waves scattered from them

SYUN-ICHI AKASOFU: On the ionospheric heating by hydromagnetic waves connected with geomagnetic micropulsations

J. S. GREENHOW, E. L. NEUFELD, and C. D. WATKINS: The scattering of 36 Mc/s radio waves by weak auroral ionization

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 255 to 270. Pergamon Press Ltd. Printed in Northern Ireland

Fading and attenuation of high-frequency radio waves propagated over long paths crossing the auroral, temperate and equatorial zones*

K. C. YEH† and O. G. VILLARD, JR.
Radio Propagation Laboratory, Stanford University,
Stanford, California

(*Received* 30 *March* 1959)

Abstract—This investigation is primarily concerned with the fading and attenuation of high-frequency radio signals propagated over a long path crossing the auroral zone. The fading of high-frequency signals propagated over non-auroral paths of comparable length has also been studied, and some new results are obtained. The principal fading and attenuation measurements on which these conclusions are based were carried out in August 1957.

For the auroral paths, there is no diurnal variation in fading speed except for a distinct minimum in time interval 1330–1900 PST, during which time the fading speed has little apparent dependence on magnetic activity along the path. In other time periods a positive correlation between magnetic activity and fading speed is found. It is suggested that the period of minimum fading speed is a consequence of the existence at that time of the kind of propagation mode made possible by ionospheric tilts.

Attenuation over the long auroral-zone path is found to be associated with "polar blackouts" as indicated by the absence of returned echo in vertical sounders located along the path. The percentage association varies with the location of the station relative to the path. This variation is consistent with the inferred propagation modes. It is found that during the hours 1330–1900 PST the attenuation cannot be attributed to the absorption that gives rise to blackouts as it can in the other hours. This is also explainable on the basis of the postulated tilt-mode propagation.

Similar observations for temperate-latitude and transequatorial paths of comparable length indicate that there is strong diurnal variation in fading speed. Some plausible explanations are offered.

OL.
17
9/60

1. INTRODUCTION

IT has been known for a long time that radio waves reflected from the ionosphere fade, and one type of fading is caused by the interference between waves propagated by two or more modes and arriving at the receiver over slightly different paths from the transmitter. However, even when arrangements have been made to receive the waves propagated by one mode the received amplitude is still found to fade. This is because there exist irregularities in the ionosphere. The wavefront of an originally plane wave becomes corrugated after reflection from such an ionosphere. Hence, an irregular diffraction pattern is produced on the ground; this pattern is affected profoundly by the processes taking place in the ionosphere.

Perhaps, one of the most intense irregularities in the ionosphere is caused by the corpuscular bombardment which simultaneously produces visual auroral display. Signals reflected directly back from such auroral ionization are characterized by rapid fading. This fading is approximately 10 to 100 times faster than that associated with ionospheric propagation in the temperate latitude (BOWLES, 1952). Frequency measurements indicate that the power spectrum of the returned echo is Doppler-shifted as well as spread out in frequency (BOWLES, 1954;

* The research reported in this paper was sponsored by the Geophysics Research Center, Air Research and Development Command, under Contract AF19(604)-1830. This paper was presented at the URSI Meeting, April 1958 at Washington, D.C.

† Now at the Department of Electrical Engineering, University of Illinois, Urbana, Illinois.

McNAMARA, 1955). The Doppler shift is approximately proportional to the operating frequency (BOOKER *et al.*, 1955).

In addition to fading, there is also absorption associated with the bombardment. The collision between the auroral particles and atmospheric atoms may generate X-rays which in turn ionize the *D*-region (CHAPMAN and LITTLE, 1957). When it is intense enough it produces what is commonly known as "polar blackout". The geographical distribution of this auroral absorption zone is found to coincide approximately with the visual auroral zone (AGY, 1954a, b). In contrast with the near midnight maximum of visual and radio auroral occurrences, auroral absorption is most intense in day-time (LITTLE and LEINBACH, 1958). The explanation is offered that even though the day-time corpuscular bombardment is weaker in strength than that occurring at night, the absorption is higher due to photodetachment of electrons at altitudes where the collision rate (and hence the absorption) is high (CHAPMAN and LITTLE, 1957; LITTLE and LEINBACH, 1958).

The present investigation is primarily concerned with studies of fading and attenuation of high-frequency radio waves propagated over different paths. The experimental set-up is given in Section (2). The statistical properties of the signals are discussed in Section (3). As a result of the conclusions of Section (3) a fading speed is defined as the average number of maxima per second. When the fading records are so analysed some interesting results are obtained. These together with the study of attenuation are given in Section (4). The interpretation of the experimental results is offered in Section (5) and it is compared with results obtained for other paths in Section (6). Finally, conclusions are offered in Section (7).

VOL.
17
1959/6

2. EXPERIMENTAL SET-UP

From 24 July to 7 September 1957, shortwave transmissions from various stations were recorded around the clock at Palo Alto, California. The principal path of interest is the one from London to Palo Alto, California, a great circle distance of 8600 km. This path traverses the auroral zone twice. For comparison purposes signals from Netherlands, Equador, Buenos Aires, Australia, Singapore and Peiping were also monitored regularly. In all cases the frequencies used were in the international broadcasting bands of 9, 11, 15 and 21 Mc/s. The 17 Mc/s band could not be used due to local interference.

The equipment used was conventional. Because of the relatively large number of stations and frequencies of interest, manual operation was employed throughout. The operator on duty chose the station from the operating schedule and recorded the carrier frequency fading on a recorder whose upper response was about 70 c/s. The signal strength was then calibrated with a signal generator.

On 15 and 21 Mc/s, rotatable three-element Yagi antennae were used. An array of twelve sloping vee antennae separated by 30° in azimuthal direction was used for recording signals of other frequencies.

3. AMPLITUDE PROBABILITY DENSITY FUNCTION AND AUTOCORRELATION FUNCTION

A large number of fading records of short-wave signals propagated over long distances, either through the auroral zone or parallel to it or across the equator, has

been analysed to find the amplitude probability density function. In almost all the cases a close approximation to Rayleigh distribution has been found. This was also found by others (GRISDALE *et al.*, 1957). An example of the record of a BBC London transmission on 15·070 Mc/s is shown in Fig. 1(a). Its amplitude distribution is found using the method of FURTH and MACDONALD (1947). It is seen that Rayleigh curve fits the experimental points quite well as shown in Fig. 1(b).

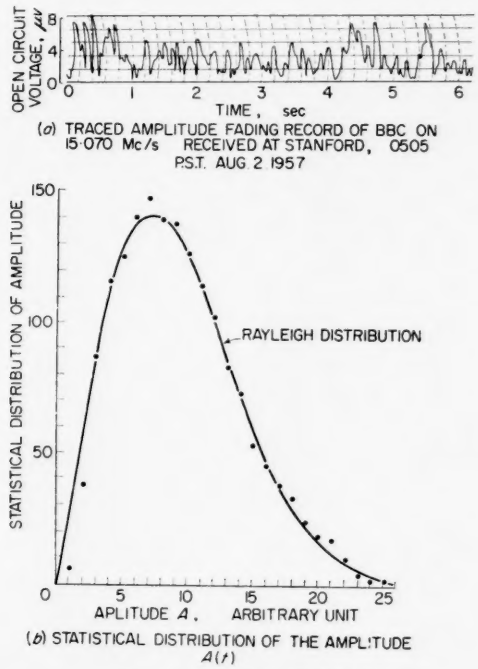


Fig. 1. Amplitude distribution.

The normalized autocorrelation function of the amplitude for the record shown in Fig. 1(a) is found by the procedure of MCNICOL (1949). As seen in Fig. 2 the Gaussian curve approximates the experimental points closely. Since the experimental information on instantaneous phase was not observed, the power spectrum cannot be found uniquely. An *a priori* assumption is necessary to check whether the assumed power spectrum is consistent with the observed amplitude autocorrelation function (BRIGGS and PHILLIPS, 1950; BOOKER *et al.*, 1950). If we assume that the power spectrum is Gaussian, then it is consistent with the observed Gaussian amplitude autocorrelation function. A large number of fading records tested have a Gaussian amplitude autocorrelation function. However, there are a small portion of the records possessing something distinctly different. A discussion of the signals of this sort is published elsewhere (YEH and VILLARD, 1958).

There are many ways of defining the fading speed of an amplitude record, but they all depend on the exact shape of the power spectrum. For simplicity we

shall define the fading speed as the average number of maxima per second. This fading speed is related to the power spectrum in a complicated manner (Rice). If the power spectrum has a Gaussian shape the fading speed is then equal to 2·52 times the frequency spread. Therefore in order to have a meaningful index of fading, the above-mentioned records which do not possess a Gaussian spectrum have been excluded from the analysis.

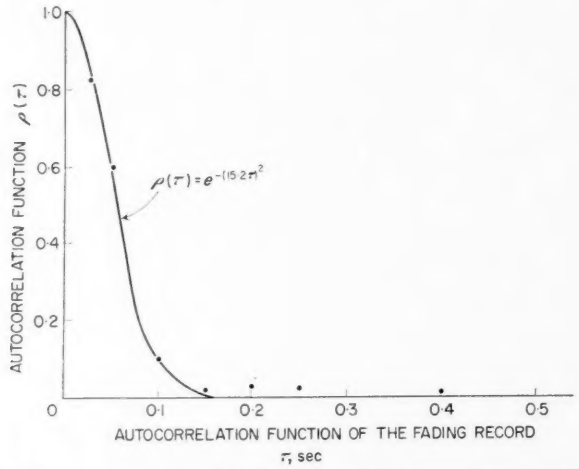


Fig. 2. Autocorrelation function of the fading record shown in Fig. 1(a).

4. Experimental Results

Almost all the short-wave transmissions were received via great-circle paths even when the transmitting antennae were directed at other parts of the world. In some cases, signals were also received in a direction corresponding to the long arc of the great-circle. Unless otherwise specified, the experimental results assembled here are for the components of signals arriving at the receiving site via the short arc of the great-circle path.

It should be noted that there are three limitations that prevent the recording of continuous data: (a) there is high D-region absorption on lower frequencies at noon; (b) there is m.u.f. (maximum useable frequency) failure at night on higher frequencies and (c) the frequencies may not be transmitted continuously through the 24 hr period.

4.1. *Auroral fading*

Signals reflected directly from the auroral ionization are characterized by rapid fading. This is demonstrated by Figs. 3(a) and (b). Fig. 3(a) is the fading record of HCJB at Quito, Equador, on 15·116 Mc/s taken on 4 September 1957. The antenna direction was in the great-circle direction. At the same time, auroral ionization forms existed in the ionosphere to the north of Palo Alto. When the Yagi antenna was rotated to this direction the fading speed was increased

VOL. 17 1959/6

by more than thirty times;* simultaneously the voice quality of the broadcasting programme became totally unintelligible.

The path that has been studied the most completely is the one from London to Palo Alto. Diurnal variation of the median fading speed on BBC 15·070 Mc/s at London is shown in Fig. 4(a).† Lines have been used to show the upper and lower decile values. As seen in Fig. 4(a), the fading speed does not have a strong diurnal variation except for a distinct minimum in the afternoon hours between about 1330–1900 PST during which time even the spread of values (as indicated

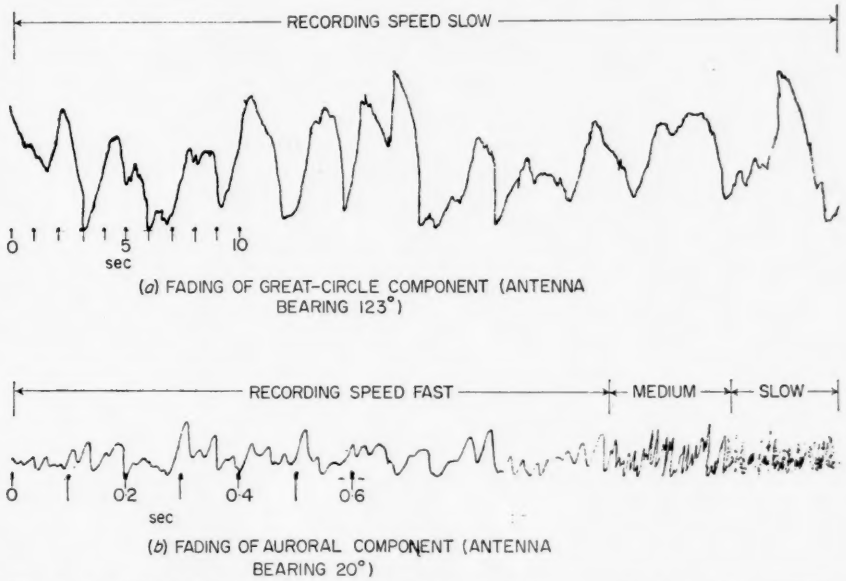


Fig. 3. Amplitude fading of great-circle and auroral components for station HCJB, Quito, Ecuador, on 15 Mc/sec, 1635 PST, 4 September 1957.

by the length of lines) becomes small. As a result the diurnal variation can be divided into three time intervals, namely 0500–1300, 1330–1900 and 2000–2400 PST. The fading is typically high and the spread of values large in both 0500–1300 and 2000–2400 PST time intervals; and is low in 1330–1900 PST.

The diurnal variation of the fading speed on other BBC frequencies and on Radio Netherlands are shown in Figs. 4(b) and (c). Note that the fading speed is not proportional to radiofrequency. Rather, it seems to be independent of it. Although such a proportional relation exists roughly for shorter path (one-hop propagation), it is generally not realized for long paths on account of the existence of many propagation modes.

The minimum observed on BBC 15·070 Mc/s again shows up on BBC 21 and 9 Mc/s and Radio Netherlands 15 Mc/s signals. The bearing and path length of

* Fig. 3(a) was taken with slow recording speed, while Fig. 3(b) with three recording speeds which increased by multiples of five.

† From 2300 to 2400 PST, data on 15·110 Mc/s were used to supplement the 15·070 Mc/s data.

Netherlands–Palo Alto path is not much different from those of London–Palo Alto path. Thus similar diurnal behaviour is expected.

The dependence of fading speed on magnetic activity along the path is illustrated in Fig. 5. Because of the distinct features in the diurnal fading variation it seems to be reasonable to divide the time of day into three intervals. The reported 3 hr magnetic K-indices for Meanook, Godhavn, and Eskdalemuir are used. The sum of the indices of these three stations is used as an indication of the magnetic activity along the path. Figs. 5(a) and (c) show that the fading speed

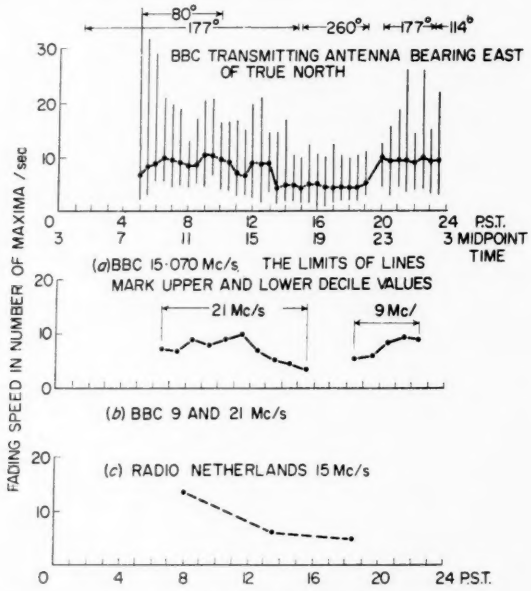


Fig. 4. Diurnal variation of the median fading speed for signals traversing the auroral zone, 24 July to 7 September 1957.

tends to be faster as magnetic activity increases. However, Fig. 5(b) shows that the fading is relatively independent of the magnetic activity in 1300–1900 PST time interval. This is somewhat unexpected. The interpretation will be attempted later in Section V.

4.2. *Auroral attenuation*

Magnetic indices at Meanook, Godhavn and Eskdalemuir observatories were used to determine the magnetically quiet days. These are 26 July, 16, 22, 23 August and 7 September, 1957. The median signal strength measured on these days is used as a reference value for different hours of the day. The reduction in signal strength obtained on other days is termed the attenuation in decibels (dB) which is equal to twenty times the logarithmic ratio of the signal strength to the reference value.

The diurnal variation of the attenuation on BBC 15·070 Mc/s is shown in Fig. 6. The median attenuation is used in (a) and the frequency of attenuation

VOL. 17 1959/6

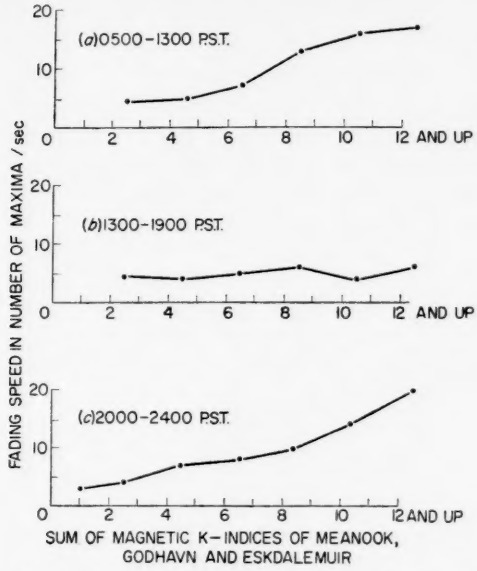


Fig. 5. Dependence of median fading speed on magnetic activity in three time intervals on BBC 15·070 Mc/sec, 24 July to 7 September 1957.

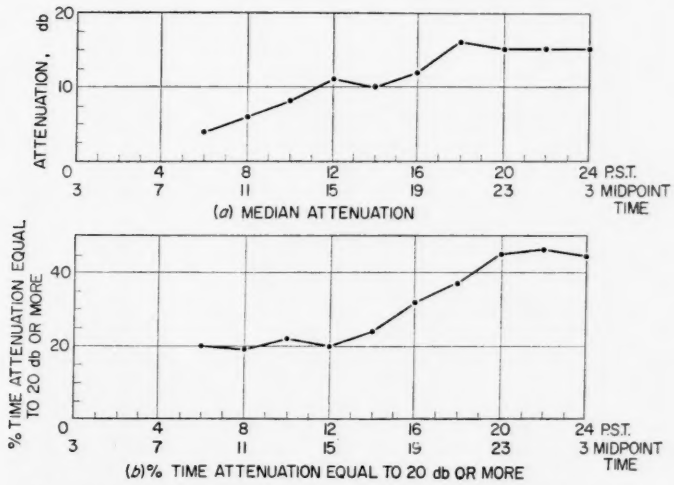


Fig. 6. Diurnal variation of attenuation on BBC 15·070 Mc/sec, 24 July to 7 September 1957.

equal to 20 dB or higher is used in (b). It is seen that both curves are similar in gross features. The maximum attenuation occurs at a time corresponding to midnight at the path midpoint. It should be noted that an investigation fails to show that the attenuation is caused by the m.u.f. failure. Rather, it seems to be associated with "polar blackouts". This is demonstrated in Table 1.

In preparing Table 1, two ratios have been used. The first one is a ratio of the

number of cases of attenuation equal to 20 dB or more when blackout (or blanketing) is observed simultaneously at the station, to the total number of cases of such high attenuation. Hence it is the maximum percentage time that high attenuation may be attributed to blackout (or blanketing). The second one is a ratio of the number of cases of high attenuation (20 dB or more) when blackout (or blanketing) is observed simultaneously, to the total number of blackouts (or blanketing). This is an indication of the maximum effectiveness of blackout (or blanketing) in producing signal attenuation.

Table 1. Attenuation on BBC 15·070 Mc/s and blackout (blanketing) occurrences on $h'f$ records (24 July to 7 September 1957)

Hours PST	Reykjavik				Churchill		Meanook	
	$A1$ (%)	$A2$ (%)	$B1$ (%)	$B2$ (%)	$B1$ (%)	$B2$ (%)	$B1$ (%)	$B2$ (%)
0500–1300	—	—	62	69	25	51	41	84
1400–1900	—	—	27	53	15	50	20	58
2000–2400	28	75	63	84	0	—	3	—

$A = f_0F^2$ blanketing by E_s; B = complete blackout; 1 = percentage association of blackout (blanketing) with attenuation; 2 = percentage association of attenuation with blackout (blanketing).

It is seen from Table 1 that more than 90 per cent of the high attenuation could be attributed to either E_s blanketing or complete blackout at Reykjavik during 2000–2400 PST. Also, during the occurrence of either event the BBC signal has a high probability (75 per cent for blanketing and 84 per cent for blackout) of being attenuated, while at the same time there is no blackout observed at Churchill and only one at Meanook. This suggests that during this time interval the attenuation is controlled quite strongly by the events happening at Reykjavik.

During 0500–1300 PST both Reykjavik and Meanook seem to have equal control over the signal strength while Churchill has only weak control. This is demonstrated when an effort to bring out the local phenomenon is made by deleting cases corresponding to simultaneous blackout in all three stations. The result shows that out of ten blackouts occurring at Churchill that did not simultaneously occur at the other two stations, only one case corresponds to high attenuation. This suggests that the signal was propagated predominately by a distinct mode. The ray path was in the D-region above Reykjavik and Meanook but not so above Churchill. In this connexion a three-hop mode of slightly unequal hops is most satisfactory to explain the geographical control of the signal strength.

In the time interval 1400–1900 PST the attenuation cannot be readily attributed to D-region absorption like polar blackout because of the low percentage values observed in $B1$ columns in Table 1 for all three stations. It should be

VOL. 17 1959/6

reminded that this is also the period of the observed low fading speed. Any interpretation of the results must simultaneously explain both phenomena.

The relation between attenuation and magnetic activity is illustrated by Fig. 7. They all show an increased attenuation at higher magnetic activity.

4.3. *Relationship between attenuation and fading*

There are a number of cases in which the attenuation and the fading speed of the BBC signals increase simultaneously. Fig. 8 is an interesting example because it shows three successive increases as indicated by the arrows. However, there

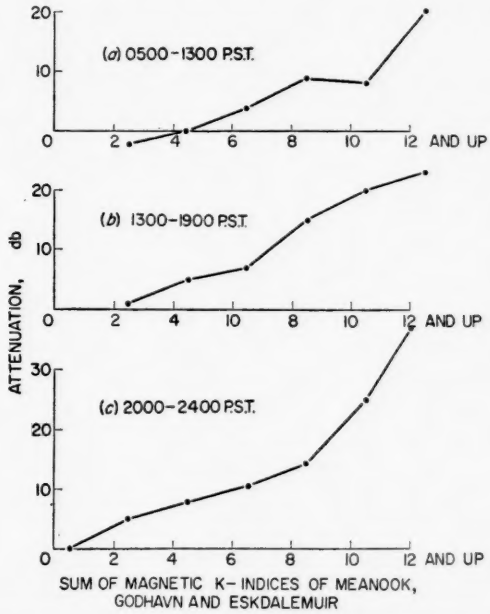


Fig. 7. Relation between median attenuation and magnetic activity in different hour intervals on BBC 15·070 Mc/s (24 July to 7 September, 1957).

are also many instances where the fading speed and attenuation do not seem to be related. Such is the case near 2200 PST in Fig. 8, when the fading speed has increased suddenly but the attenuation does not change greatly.

Cases that show the relation between attenuation and fading are not numerous enough to enable a statistical study. It can only be said that cases of both kinds were observed without apparent preference of the time of the day.

4.4. *Component propagated along the long-arc of the great circle*

BBC signals propagated along the long-arc of the great-circle path were identified and received consistently from 0500 to 1000 PST on 15 and 21 Mc/s. The path length is 31,400 km, more than three times that of the direct path. It passes through the southern auroral zone. The signal characteristics are almost identical to the direct signal in voice quality and fading. Within these 5 hr the

median fading speed on both 15 and 21 Mc/s is roughly constant at about 10 max/sec, although it varies considerably in a single day. The signal strength is lower than that of the direct path but may occasionally rise to a value above it. One point of interest is that the predicted m.u.f. near the southern auroral zone is lower than 15 Mc/s and much lower than 21 Mc/s. It is highly probable that the transmission was via propagation modes associated with ionospheric tilts (VILLARD *et al.*, 1957; STEIN, 1958).

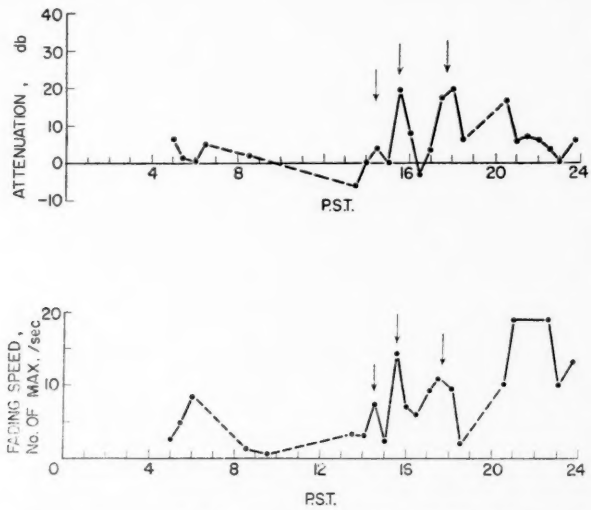


Fig. 8. Attenuation and fading speed on BBC 15·070 Mc/s, 8 August 1957 (arrows indicate simultaneous rises and falls).

5. INTERPRETATION

The diurnal variation of the fading speed is shown in Fig. 4. The fading is high in the time intervals 0500–1300 PST and 2000–2400 PST. During these time intervals the fading speed increases with an increase of magnetic activity along the path. This is expected because of the known close relation between auroral occurrence and magnetic activity. Hence, the fading speed seems to be a rough index of the magnetic activity along the path.

The result of a study of attenuation is shown in Table 1 and Fig. 7. It is seen during these two time intervals the attenuation could be attributed to the *D*-region absorption that gives rise to the polar blackouts, and there is increased attenuation with magnetic activities. The study by AGY (1954a, b) shows that during magnetic activity the quiet day auroral absorption belt around the pole expands out by several folds in latitude from its normal 6° width. Also, the absorption of the extraterrestrial radio waves on 30 Mc/s has a positive correlation with magnetic *K*-index (LITTLE and LEINBACH, 1958). Therefore, it seems reasonable to expect such relation to be carried over to the attenuation over an oblique path. Additional attenuation may arise from scattering of waves from field-aligned ionization columns in the *E*- and *F*-regions, which are presumably produced by particle bombardment. This attenuation is expected to be most pronounced when the

VOL. 17 1959/6

wave is propagating nearly perpendicular to these columns as is the case for high latitude oblique propagation.

Recent auroral radar results reveal that auroral radar echo occurrence has the usual night-time maximum outside the auroral zone but a midday maximum inside the auroral zone (McNamara, 1958). The diurnal behaviour of the fading might be explained as the superposition of these two maxima since the path traverses from outside of the auroral zone to well inside of the zone and then out of the zone again. Undoubtedly the fading is affected by such a phenomenon.

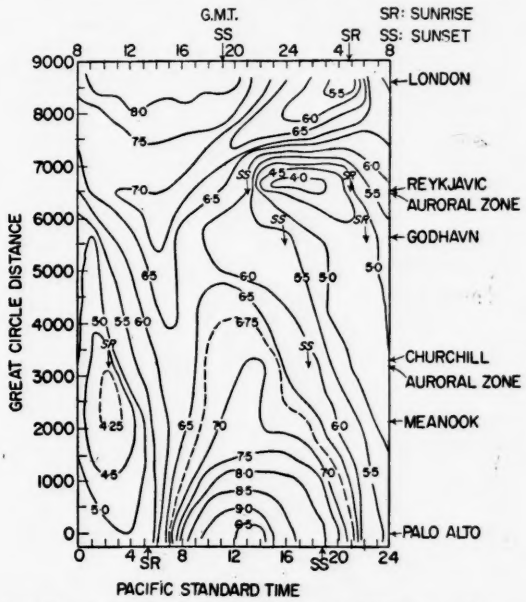


Fig. 9. Contours of median f_0F^2 values (Mc/s) for August 1957 along the great-circle path from Palo Alto, California to London.

But, while the superposition hypothesis may explain the minimum in fading in 1330–1900 PST, it seems to meet difficulties in explaining why, during the same time interval, (a) the fading is relatively independent of the magnetic activity shown in Fig. 5(b), and (b) the attenuation of the signal cannot be readily attributed to D-region absorption shown by the low percentage values in $B1$ columns in Table 1 while it is correlated with magnetic activity shown in Fig. 7(b). It seems that perhaps other explanation is necessary.

In order to investigate this further, contours of median f_0F^2 are plotted in Fig. 9 for August 1957. These contours are constructed from the reported vertical-incidence data (CRPL, 1958). The period of interest is 1330–1900 PST or 2130–0300 GMT. Let us temporarily focus our attention on f_0F^2 values near the Reykjavik station. It is seen that 1300 PST corresponds to sunset time at Reykjavik. The critical frequency falls off very quickly as it becomes dark at this station. This makes possible a strong gradient of ionization density (or ionospheric tilt). In the meantime a gradient of opposite sign exists at Palo Alto owing to increasing

obliquity of the sun's rays towards the pole. Further, these opposite gradients exist near the transmitting and receiving control points only in time 1330–1900 PST. Although the critical frequencies do not give a full information of the ionization density profiles, they are approximate and fair representations of them. It is felt that the gradients are large enough to have a predominant effect on the propagation mode, i.e. to excite the propagation modes that are supported by ionospheric tilts (VILLARD *et al.*, 1957; STEIN, 1958). When a tilt of the order of a few degrees exists, a reflected ray may miss the earth entirely and illuminate the ionosphere at a second reflection point. Such reflections can go on until a tilt of

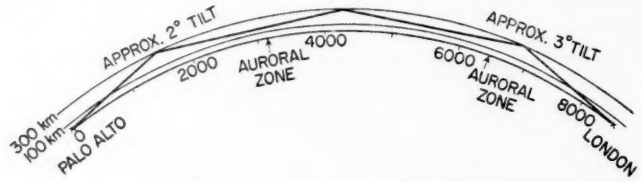


Fig. 10. Mode of propagation supported by ionospheric tilts (^3F).

opposite sign is met and reflects the ray back to the ground. The tilts are estimated to be roughly 2° and −3° at the two control points. Based on geometrical considerations the propagation between London and Palo Alto in the period 1330–1900 PST is most likely via a ^3F-mode. The superscript of F designates the number of successive ionospheric reflections. Fig. 10 illustrates such a mode.

From Fig. 10 it is seen that, except near two end points which are outside of the auroral zone, the ray always stays above the E-region, the height of maximum auroral occurrence. For this reason the fading speed will be relatively less affected by auroral occurrences. It can be postulated that during magnetically quiet days there are components propagated by conventional modes in addition to this tilt-supported mode during these hours. Consequently, these different components will beat and produce higher than normal fading. During moderate magnetic disturbances the component propagated by conventional modes may become gradually attenuated resulting in a reduced total signal strength. At the same time the fading of the tilt-supported mode may increase somewhat owing to the formation of field-aligned ionization irregularities in the E- and F-regions. It is possible that this slight increase in fading speed is approximately counterbalanced by the fact that there is less multiple-mode beating, owing to the disappearance of conventional modes. Hence, for this path, the fading speed may remain approximately unchanged at times of higher magnetic activity. On the other hand, attenuation is somewhat increased by the magnetic activity. An additional source of attenuation which may also be important, is considered in the following.

There is evidence at present that the ionospheric wind velocity is increased in the horizontal direction during period of magnetic activity, particularly for the F-region (BRIGGS and SPENCER, 1954). While the precise variation of the profile of the layer during magnetic activity is not known, it may be reasonably assumed that during such a period there is perturbation to the otherwise smooth and steady layer. The propagation of waves supported by the ionospheric tilt is expected to

VOL.
17
1959/6

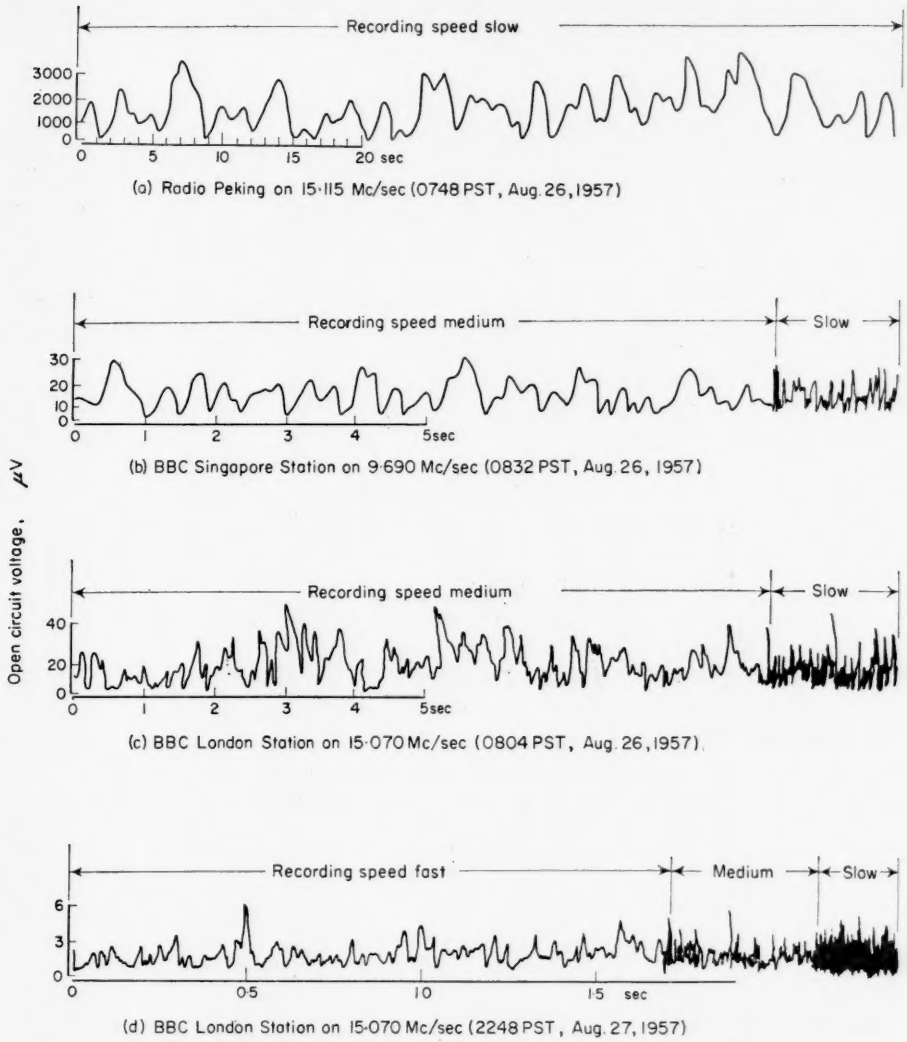


Fig. 11. Comparison of fading records for signals propagated over different geographical zones.

be very sensitive to slight changes in electron density contours. The tilt-mode may even become unstable at times. As a result the signal strength received during such occasions will be reduced.

As indicated in Section (4.3) the fading speed and the attenuation may or may not increase simultaneously. LITTLE and LEINBACH's work (1958) indicates that particle bombardment produces absorption layer in day-time but not at night. Therefore we may expect simultaneous rise in attenuation and fading in day-time and a rise in fading, but not in attenuation, at night. While there are not enough cases observed to make a statistical study, no such day-time and

night-time phenomena have been found. It is felt that, since we are dealing with an oblique path, fading and attenuation of the wave may have been introduced at different localities on the path.

6. COMPARISON WITH OTHER PATHS

The fading of radio waves has been studied by MCNICOL (1949) for short paths and by GRISDALE *et al.* (1957) for long paths. Except for the case of nearly vertical incidence the amplitude of the fading record has a Rayleigh distribution and the power spectrum is consistent with a Gaussian distribution.

Shown in Fig. 11 are some typical examples of the fading speed of various paths arranged in an order of increasing fading speed. The fading of Radio

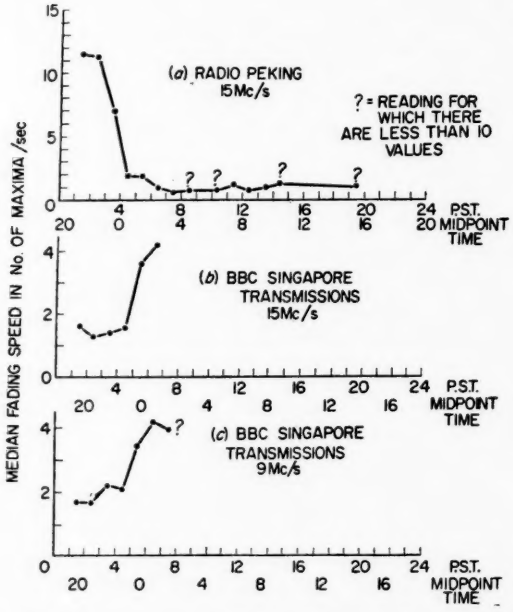

MEDIAN FADING SPEED IN No. OF MAXIMA /sec
(a) RADIO PEKING 15Mc/s
? = READING FOR WHICH THERE ARE LESS THAN 10 VALUES
(b) BBC SINGAPORE TRANSMISSIONS 15Mc/s
(c) BBC SINGAPORE TRANSMISSIONS 9Mc/s
P.S.T. MIDPOINT TIME

Fig. 12. Diurnal variation of the median fading speed for Radio Peking and BBC Singapore Transmissions (August 1957).

Peking is typical for a temperate-latitude path and it is slowest. The trans-equatorial paths possess intermediate fading speeds between temperate-latitude paths and auroral paths, and this is demonstrated by the example shown in Fig. 11(b). The fading is in general highest for the auroral paths as shown in Fig. 11(c), and the auroral fading may become as violent as that of Fig. 11(d).

The preliminary results obtained indicate the existence of diurnal variations of the fading speed. These are shown in Fig. 12. At this time we cannot draw firm conclusions as to the cause of diurnal variations. It is speculated that the peak in fading of Radio Peking is actually caused by auroral disturbances when

VOL. 17 1959/6

they move southward (MITRA, 1952) into the path. As for the increase in fading of BBC Singapore transmission, it is possible that the equatorial spread-F may play a role. At Singapore spread-F activity reaches a maximum at about 2100 local time (OSBORNE, 1951). The time of high fading corresponds closely to the period of high spread-F activity near Singapore.

7. CONCLUSIONS

A large number of amplitude fading records have been obtained from 24 July to 7 September 1957. These records are for signals propagated over a long path (6200–14,700 km) through different geographical zones. Most of these signals were received in a great-circle direction even when the main bearing of the transmitting antennae was directed elsewhere. For the few examples tested the statistical properties of the amplitude distribution and autocorrelation are similar in shape. Signals obviously possessing different properties have not been included in the present analysis. When data taken continuously are analysed, the following results are obtained.

7.1. *Auroral path*

For the London to Palo Alto path:

(a) There is no diurnal variation in fading speed except for a distinct minimum from 1330–1900 PST.
(b) In the hours 0500–1300 and 2000–2400 PST, fading speed increases with magnetic activity.
(c) In the hours 1330–1900 PST, the fading is relatively independent of magnetic activity.
(d) Attenuation increases with higher magnetic activity in all hours of the day.
(e) When a blackout is observed at a vertical sounder along the path, high attenuation over the oblique path occurs simultaneously 50 per cent of the time or more.
(f) In the hours 0500–1300 and 2000–2400 PST high attenuation can be attributed to the absorption that gives rise to the polar blackout.
(g) In the hours 1330–1900 PST high attenuation cannot be readily attributed to the absorption that gives rise to the polar blackout.

The present investigation has demonstrated the close relation between fading and attenuation over the oblique path on the one hand, and the magnetic activity and blackout observed locally along the path on the other hand. Because of the close relation between auroral activity and magnetic activity, a positive correlation of fading speed and magnetic activity might be expected. This is found to be true except for the hours from 1330 to 1900 PST. The deviation from expectation observed in these hours is explained as due to change in the mode of propagation. Owing to the presence of strong gradients of ionization which are present only in these hours, a new mode is excited by ionospheric tilts. With such a mode the portion of the ray in the auroral zone stays above the 100 km level at which the maximum auroral activity occurs. In such a case, the fading will be low and above all it will be relatively independent of the magnetic activity.

7.2. *Non-auroral paths*

It has been found that the average fading speed of a radio signal depends very much on the paths. In the order of increasing fading speed these paths are temperate-latitude path, transequatorial path and auroral path.

The results on non-auroral paths are not as complete as auroral paths. Nevertheless, they exhibit strong diurnal variations. There is some evidence as to the cause of high fading for these paths. The indication is that the auroral-like fading of Radio Peking was actually caused by the aurorae when they moved southward into the path. Also, the high fading for Singapore–Palo Alto path suggests the connexion between fading and equatorial spread-*F*. However, more evidence is required before the cause of fading is finally settled.

Acknowledgements—The writers wish gratefully to acknowledge the co-operation and assistance given to them from various sources. Messrs. F. AXON and P. E. F. A. WEST of the British Broadcasting Corporation have provided information on their transmissions. Magnetic data were supplied by Mr. R. G. MADILL, Dominion Observatory, Ottawa, Canada, by Mr. V. LAURSEN, Det Danske Meterologiske Institut, Charlottenlund, Denmark, and by Dr. J. VEDLKAMP, Kon. Nederl. Meterorl. Institut, De Bilt, Holland.

The operators were D. WESTOVER, G. FAN, K. RENSHAW and S. SAPHE, all students at Stanford University. The long and tedious task of record analysis was mostly Mrs. R. S. RICH's effort. She was also a student at Stanford University.

VOL.
17
1959/6

REFERENCES

AGY V.	1954a	*J. Geophys. Res.* **59,** 267.
AGY V.	1954b	*J. Geophys. Res.* **59,** 499.
BOOKER H. G., GARTLEIN C. W. and NICHOLS B.	1955	*J. Geophys. Res.* **60,** 1.
BOOKER H. G., RATCLIFFE J. A. and SHINN D. H.	1950	*Phil. Trans.* A **242,** 579.
BOWLES K.	1952	*J. Geophys. Res.* **57,** 191.
BOWLES K.	1954	*J. Geophys. Res.* **59,** 553.
BRIGGS B. H. and PHILLIPS G. J.	1950	*Proc. Phys. Soc.* (B) **63,** 907.
BRIGGS B. H. and SPENCER M.	1954	*Report on Progress in Physics* Vol. XVII, p. 245. Physical Society.
CHAPMAN S. and LITTLE C. G.	1957	*J. Atmosph. Terr. Phys.* **10,** 20.
CRPL	1958	*Ionospheric Data* Pt. A.
FURTH R. and MACDONALD D. K. C.	1947	*Proc. Phys. Soc.* **59,** 388.
GRISDALE G. L., MORRIS J. G. and PALMER D. S.	1957	*Proc. Instn. Elect. Engrs.* B **104,** 39.
LITTLE C. G. and LEINBACH H.	1958	*Proc. Inst. Radio Engrs., N.Y.* **46,** 334.
MCNAMARA A. G.	1955	*J. Geophys. Res.* **60,** 257.
MCNAMARA A. G.	1958	Private communication.
MCNICOL R. W. E.	1949	*Proc. Instn. Elect. Engrs.* Pt. 3, **96,** 517.
MITRA S. K.	1952	*The Upper Atmosphere* p. 417. The Asiatic Society, Calcutta.
OSBORNE B. W.	1951	*J. Atmosph. Terr. Phys.* **2,** 66.
RICE S. O.	1954	*Noise and Stochastic Processes* (Edited by WAX N.). Dover Publications, New York.
STEIN S.	1958	*J. Geophys. Res.* **63,** 217.
VILLARD O. G., STEIN S. and YEH K. C.	1957	*J. Geophys. Res.* **62,** 399.
YEH K. C. and VILLARD O. G.	1958	*Proc. Inst. Radio Engrs., N.Y.* **46,** 1958.

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 271 to 275. Pergamon Press Ltd. Printed in Northern Ireland

A note on fourth reflection condition in the ionosphere

B. CHATTERJEE
Indian Institute of Technology Kharagpur, India

(*Received* 21 *June* 1959)

ACCORDING to the Appleton–Hartree magneto-ionic formula for wave propagation at vertical incidence in the ionosphere, the incident wave is split up into two components—ordinary and extraordinary—due to presence of earth's magnetic field (APPLETON 1932). For the case of transverse (i.e. magnetic field perpendicular to the direction of wave travel) or quasi-transverse propagation and for wave-frequency (f) greater than the gyro-frequency (f_h), the refractive index (μ) for the extraordinary wave for a region of given ionization density N (or critical frequency f_0) will be zero for two values of f and infinity for another value. All these three values of f are being determined by f_0 and f_h.

The condition for which μ is infinity is often referred to as the "fourth condition of reflection" (MITRA, 1952). The group velocity of a wave packet tends to zero as the refractive index tends to zero *or* infinity. As such it was thought before that the wave packet will be reflected back at vertical incidence also for the case $\mu = \infty$, just as in the case $\mu = 0$. Some observers (PANT and BAJPAI, 1937, MARTYN and MUNRO, 1938) even reported the existence of such a reflected wave. But HEADING and WHIPPLE (1952) in their analysis found the anomaly that for the fourth reflection condition there is appreciable absorption even when the collisional frequency tends to zero, and doubts were raised regarding the existence of such a reflected wave. BUDDEN (1954) has shown that a wave packet cannot be reflected from a region having infinite refractive index. He concluded from his analysis that, although the extraordinary wave approaches the level $\mu = \infty$, it never reaches the same but moves horizontally in the magnetic meridian. Thus it never gets reflected.

The same phenomenon of no reflection at $\mu = \infty$ can also be simply explained by "ray-treatment" if we plot μ (or μ^2) against f (or f^2), as done in Fig. 1 below, and not μ against N or f_0 as is normally done. The latter type of plotting is, of course, very useful for understanding the propagation of a wave packet inside an ionospheric layer, but it does not explain the phenomenon taking place at the region for which $\mu = \infty$ for the given wave.

Fig. 1 shows the variation of μ^2 against the square of the normalized value of wave-frequency [i.e. $(f/f_0)^2$] for the transverse propagation of extraordinary wave at vertical incidence in a given homogeneous ionized layer (i.e. for a given value of N or f_0). It is assumed that $f_H = 0{\cdot}2\,f_0$ and the plotting is done for the case $f > f_H$ i.e. for the case where the so-called "fourth reflection condition" exists.

An ionized region can be visualized as made up of a large number of horizontally stratified adjacent layers of different ionization densities as shown in Fig. 2.

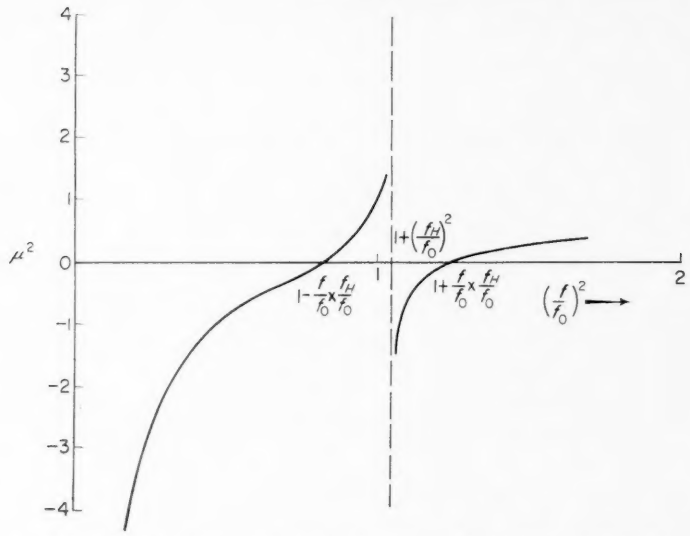
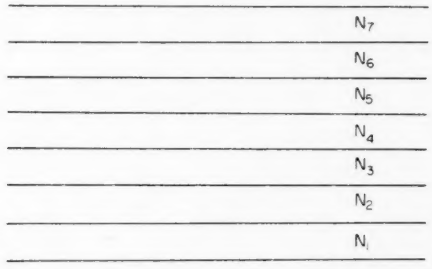


Fig. 1. Variation of μ^2 with $(f/f_0)^2$ for the transverse propagation of extraordinary wave in a region of given ionization density.



$N_1 < N_2 < N_3 < N_4 < N_5 < N_6 < N_7 \ldots\ldots$

Fig. 2. An ionized region with ionization density increasing, with height being represented as a large number of adjacent homogeneous layers of different ionization densities.

Provided each such layer is assumed to be of very small thickness they may be considered as homogeneous, having uniform ionization density inside each layer and differing slightly from the adjacent layers. Fig. 1 shows the variation of μ^2 against $(f/f_0)^2$ for any one of such homogeneous layers (say, layer N_5 in Fig. 2).

It may be mentioned that in considering the phenomenon of refraction (at oblique incidence) from a denser to a rarer medium (i.e. from a medium of higher refractive index to that of a lower refractive index) reflection is said to occur when the sine of the angle of incidence (i) is equal to (or greater than) the relative refractive index of the second medium. The smaller the angle of incidence, the less should be the refractive index of the second medium in order to have "reflection". The case of vertical incidence is considered as a limiting case when the angle of incidence tends to zero—thus requiring a zero refractive index for the second medium for reflection to occur. It may be pointed out here that in all such cases, reflection takes place

at the boundary of the two media. Thus we see that, if the relative refractive index is less than or equal to the sine of the angle of incidence, the ray (or wave) is reflected back at the boundary and does not enter the second medium. If it is larger than sin i but less than unity, the wave enters the second medium but will be deviated away from the normal. For relative refractive index larger than unity, the ray will enter the second medium and will be deviated towards the normal—the deviation being larger, the larger the refractive index. For an infinitely larger refractive index value, the wave will move along the normal inside the second medium for any value of angle of incidence at the boundary. These facts of geometrical optics have been taken help of in explaining the phenomenon at "fourth reflection condition" in the ionosphere.

Explanation with ray treatment becomes simplified if we consider vertical incidence as the limiting case of oblique incidence, i.e. the angle of incidence tending towards zero (and not exactly zero), because, only then can we apply the theories of geometrical optics which considers the bending of rays (towards or away from the normal) at refraction. The following paragraphs consider the case of vertical incidence with this end in view (i.e. as a limiting case of oblique incidence).

It is seen from Fig. 1 that, for small values of $(f/f_0)^2$, μ^2 is negative, i.e. μ is imaginary. This means that in the layer concerned, waves of such frequencies cannot exist. If excited, they will die down within a short distance inside the layer like evanescent modes in a wave-guide. The waves get turned back or reflected from the boundary of such a layer and cannot proceed into it. Provided the collision frequency is not high, almost the whole of wave energy gets reflected (or reradiated) backwards. This situation exists till the wave frequency satisfies the relation

$$f^2 = f_0^2 - ff_h \left[\text{or } (f/f_0)^2 = 1 - (f/f_0)\left(\frac{f_h}{f_0}\right) \right].$$

For this frequency $\mu = 0$ and the extraordinary wave is just reflected back from the layer boundary. With further increase in wave-frequency, μ^2 becomes positive and increases in magnitude. As a result, the wave deviates less and less away from the normal and is able to penetrate the boundary and move into the layer concerned. For $f^2 = f_0^2$ [i.e. for $(f/f_0)^2 = 1$], the value of μ for the extraordinary wave is unity and the wave should penetrate the layer undeviated. With further increase in wave frequency, $\mu > 1$, and its magnitude increases rapidly—attaining an infinite value for $f^2 = f_0^2 + f_h^2$ [i.e. for $(f/f_0)^2 = 1 + (f_h/f_0)^2$]. For $\mu > 1$, the wave bends towards the normal and the bending is more, the higher the value of μ. As μ approaches infinity, the wave tends to move upwards along the normal. The group velocity approaches zero as μ approaches infinity and the extraordinary wave under this condition—moving infinitely slowly upwards along the normal—gets trapped in the layer concerned and cannot come back or be reflected. Thus, there cannot be any reflection for the condition $\mu = \infty$ and the wave energy gets dissipated in the region concerned (*vide infra.*). Thus, no reflection takes place for

$$(f_0^2 - ff_H) < f^2 < (f_0^2 + f_h^2).$$

It can be shown that the total absorption suffered by a radio wave in a layer of homogeneous ionization density is proportional, both to ν (the collisional frequency in the layer) and to $\int dh/v_g$, where v_g is the group velocity of the particular wave in the layer concerned and the integration is carried out over the whole thickness of the layer. Now, $\int dh/v_g$ is actually the time (t) for which the wave remains in the layer concerned and we see that the absorption suffered is not only proportional to ν but also to t. The extraordinary wave under transverse propagation and for $\mu = \infty$, moves *into* the layer concerned at an infinitely slow rate and thus remains there for an indefinitely long time. As such even for a very small value of ν, the total absorption suffered will be infinitely large and the whole energy will thus be "absorbed" or dissipated in the layer concerned.

The situation is completely different for the case $\mu = 0$. In that case, as the wave gets reflected at the boundary of the layer and does not enter into it, there cannot be any absorption in the given layer under consideration. Of course, large absorption takes place in adjoining layers below, whose μ value is just above zero—thus having low group velocity. This accounts for the large deviative absorption observed for frequencies very near the penetration frequency of an ionized layer.

From the above discussions we see that, for transverse propagation at vertical incidence in a given layer of specified ionization density (i.e. for given values of N or f_0), the extraordinary wave is not reflected for values of f^2 lying between $(f_0^2 - ff_h)$ and $(f_0^2 + f_h^2)$. But it may be mentioned here that, in normal ionospheric propagation, there always exists a layer at a lower height, having a lower ionization density or critical frequency (say $f_0' < f_0$), for which the said wave-frequency (f) will satisfy the relation

$$(f'_0{}^2 + f_h^2) < f^2 < (f'_0{}^2 + ff_h).$$

It has been shown in the next paragraph that, for waves satisfying this relation, there will be reflection from the layer concerned (of critical frequency f_0'). So waves of such frequencies in normal ionospheric propagation will be reflected from the lower layer of critical frequency f_0' before reaching the upper layer of critical frequency f_0.

With further increase in wave frequency—when $f^2 > (f_0^2 + f_h^2)$—the value of μ^2 again becomes negative, i.e. μ becomes imaginary (see Fig. 1). For such frequencies, therefore, reflection of the extraordinary wave will take place, as discussed earlier. This continues till $f^2 = f_0^2 + ff_h$, when the value of μ again becomes zero and the wave is just reflected from the boundary of the layer concerned. With further increase in f, μ^2 becomes positive (i.e. μ is real) and it approaches the value of unity asymptotically. This means that for $f^2 > (f_0^2 + ff_h)$ no further reflection occurs from the layer concerned and the wave is deviated (away from the normal) less and less as the wave frequency increases more and more. The advantage of plotting μ^2 against $(f/f_0)^2$ thus lies in the fact that the phenomenon at $\mu = \infty$ (as well as for other values of μ) can be simply explained with geometrical optics, without going into the complicated wave theory.

It is evident from Fig. 1, that the variation of refractive index with frequency

VOL
17
1959/

in a given ionized region is analogous to that of any other dispersive medium—the extraordinary wave having an "absorption band" for the frequency (say f_a) = $\sqrt{f_0^2 + f_h^2}$. As discussed earlier, the value of μ becomes infinity in such a case and the wave moves normally into the layer at an infinitely slow rate. Thus the wave remains "trapped" inside the given layer for an indefinitely long time (t) and any finite rate of dissipation (finite value of ν)—however small—will be able to dissipate the whole energy inside the layer. As the value of ν in a material medium can never be zero and the value of t is infinitely large in this particular case, the absorption of wave energy (which is proportional to νt) is also infinitely large. As a result, the whole of extraordinary wave energy under this condition gets dissipated in the layer and no reflection of wave energy can take place. This accounts for the so-called "absorption band" at this frequency.

Acknowledgements—I express my grateful thanks to Prof. H. Rakshit, D.Sc., F.N.I., F.Inst.P., Head of the Department of Electronics and Electrical Communication Engineering, for his helpful suggestions and kind interest in the work.

References

Appleton E. V.	1932	*J. Inst. Elect. Engrs.* **71,** 642.
Budden K. G.	1954	*Physics of the Ionosphere* p. 320. Physical Society, London.
Heading J. and Whipple R. J. P.	1952	*Phil. Trans. Roy. Soc.* A **244,** 469.
Martyn D. F. and Munro G. H.	1938	*Nature, Lond.* **141,** 159.
Mitra S. K.	1952	*The Upper Atmosphere* p. 192.
Pant B. D. and Bajpai R. R.	1937	*Science and Culture* **2,** 409.

VOL. 17 9/60

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 276 to 285. Pergamon Press Ltd. Printed in Northern Ireland

Photometric observations of the twilight glow [OI] 5577 and [OI] 6300

L. R. Megill
National Bureau of Standards, Boulder, Colorado

(*Received* 30 *May* 1959)

Abstract—The enhancement of both the [OI] 5577 and the [OI] 6300 atmospheric emission lines has been measured in detail for the night of 6/7 January 1958 at Rapid City, South Dakota. The enhancement of the 5577 emission cannot be explained on the basis of photo excitation alone. The 6300 emission is compared with recent theory. Fair agreement is noted between theory and observation.

The twilight enhancement of [OI] 6300 is well known; on the other hand, the twilight enhancement of [OI] 5577 has never been reported in detail. A brief reference was made by Dufay and Dufay (1948) in which they reported an enhancement of about a factor of 2. A spectroscope was used to make the measurements. Koomen *et al.* (1955) and Manring and Petit (1955) have reported no enhancement of [OI] 5577 in the twilight. We have been able to detect an enhancement of [OI] 5577 at both our station in Rapid City, South Dakota, and our station at Fritz Peak, Colorado. Photometric observations of both [OI] 5577 and [OI] 6300 made at the IGY Airglow station in Rapid City, South Dakota, on the night of 6/7 January 1958 have been analysed in detail for both colours.

VOL.
17
1959/6

Instrumentation

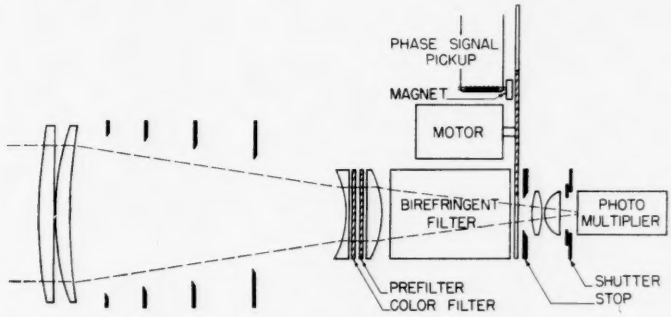
The instrument used for the observations is one which was developed at the National Bureau of Standards for the airglow programme of the International Geophysical Year. The basic component of the instrument which enables one to observe during twilight is a birefringent filter. This filter has the rather unique property that line sources are transmitted as a fluctuating light signal while any continuous background light is essentially constant. The instrument was first developed by Blamont and Kastler (1951) and since has been used by Koomen *et al.* (1955) and Manring and Petit (1955). The design for our filter is due to Dunn and Manring (1955).

The photometer used is shown schematically in Fig. 1. The prefilters used are dielectric filters with a width at half transmission of about 20 Å. Coloured glass filters are used to cut out the side bands. The field of view of the photometer is 4°.

An R.C.A. 6217 photomultiplier is used as a detector. The signal from the photomultiplier is amplified in a conventional R.C. coupled amplifier whose frequency response is peaked at 80 c/s. This is the chopping frequency of the rotating polaroid used in the birefringent filter. The output of the amplifier is fed into a phase sensitive rectifier. The reference voltage for the rectifier is derived from the same shaft as that which turns the polaroid. The signal is recorded on a strip chart by a rectilinear recording, galvanometer type recorder.

Koomen *et al.* (1955) have made extensive tests with an instrument of similar

design. When exposing this instrument to a white light source, they discovered a small percentage of ripple present due to the fact that the combination of their pre-filter and phototube response did not present a true continuum to the birefringent filter. They determined that the principal cause of the ripple component was the slope of the spectral curve of the photomultiplier used as a detector. In order to ensure that the enhancement observed was not due to instrumental causes, a white light source (in this case an incandescent lamp) was placed in front of our telescope and the effect measured. The total white light intensity can be measured



TELESCOPE OPTICAL SYSTEM

Fig. 1. This figure shows schematically the layout of the photometer telescope. The relay coil and magnet system shown is used to derive a reference voltage needed for phase sensitive rectification.

by measuring the d.c. component of the current from the phototube. In our experiment the net deflexion due to continuum chopping in a light source of the same intensity as the scattered background was negligible compared to the 5577 Å signal. An additional fact supporting the observation is that on some occasions the enhancement is not present when measured with an identical instrumental setup at the same light background intensity.

The absolute calibration of the instrument is a difficult problem. In this case, the calibration was obtained by a comparison with our instrument at Fritz Peak, Colorado, which has been calibrated by using the deflexion due to stars. The calibration of the instrument has recently been compared with other airglow photometers around the world using a portable photometer for the inter-calibration (Roach, 1958). The results all agree within about 25 per cent. It is believed that the calibration for λ5577 is accurate for about 25 per cent and about 50 per cent for the λ6300.

The Observational Data

Observations for the night of 6/7 January 1958 were taken in two ways. During the morning and evening twilights data were taken in an almucantar at 75° zenith distance. The photometer started in the north, swept around through the east, south, west and back to the north again. On the next sweep the order was reversed. Between each sweep the prefilter would be changed so the λ5577 and

λ6300 radiations were alternately observed. The sweeping rate was 10° per second. The interval between sweeps of one colour was 3 min. During the bulk of the night, when the intensities were not varying so rapidly, the standard IGY observing schedule was used. This consists of a series of almucantar sweeps at zenith distances of 80°, 75°, 70°, 60°, 40° and 0°. One such survey was obtained for each of the two colours at 15 min intervals

A typical record is shown in Fig. 2. This record shows the enhancement in both the λ6300 and the λ5577 lines in the twilight.

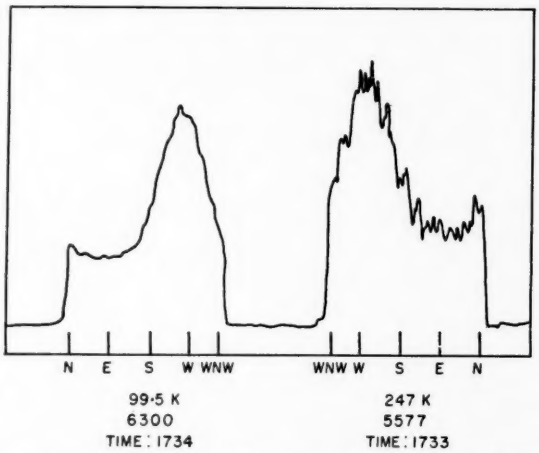


Fig. 2. The type of record obtained from the recorder is shown here. Both traces show the twilight enhancement. Telescope directions and sensitivities in kilohms (k), as well as the colour and time, are marked in below the tracings.

VOL.
17
1959/6

Analysis of the Data

The records were read at 22·5° intervals in azimuth. The intensities were then reduced to slant intensities outside the atmosphere using the relation:

$$I = QI_1[1 - Sc(75)I_{\text{avg}}]e^{\tau m}(75)$$

where I is the intensity outside the atmosphere, I_1 is the recorded intensity inside the atmosphere, I_{avg} is the "average" intensity of the sky as a whole and $Sc(75)$ is the fractional intensity scattered into the line of sight from parts of the sky not being observed, τ is the absorption coefficient per unit air mass, $m(75)$ is the air mass in the line of sight and Q is the calibration coefficient for the instrument. The values used in the reduction are given in Table 1.

I_{avg} is determined by taking an average of the sixteen readings from a given sweep and correcting to the zenith inside the atmosphere. This gives an approximate correction for the scattered light intensity which is sufficiently accurate for the work at hand. The complete expression for I_{avg} is

$$I_{\text{avg}} = \frac{\sum_1^{16} I_i(75)}{16}\left[1 - \frac{e^{\tau m}(75)}{V(75)}Sc(75)\right][1 + Sc(0)]\frac{e^{\tau m}(75)e^{-\tau m}(0)}{V(75)}.$$

The scattering coefficients, Sc, are taken from ASHBURN (1954). $V(z)$ is the Van Rijhn correction to take into account the fact that we are looking at an angle through an emitting layer.

The data are now in the form of slant intensities outside the atmosphere. In order to obtain the equivalent zenith intensity one needs but to divide by the Van Rijhn factor to correct for the fact that we are looking at an angle through the layer.

Table 1. Constants used for reduction of data to outside the atmosphere

	5577	6300
τ	0·150	0·103
$Sc(75)$	0·260	0·227
$Sc(0)$	0·067	0·059

At this point it is necessary to separate the discussion of the two emissions. The λ5577 radiation will be discussed first.

In order properly to evaluate the measurements made, it is necessary to know the position of the sun's rays in the atmosphere at the time of observation. The relationship of principal interest is height above the earth's surface at which the sun's rays and the line of sight of the photometer intersect. Fig. 3 shows the geometry of the problem. The height is first solved excluding consideration of the screening by the lower atmosphere. From the diagram the following relationships may be obtained:

$$\sin\theta = \sin\psi/(1 + z/R) \tag{1}$$

$$\gamma = \psi - \theta \tag{2}$$

$$\sin\beta = \cos\gamma \sin\alpha - \sin\gamma \cos\alpha \cos\Delta A \tag{3}$$

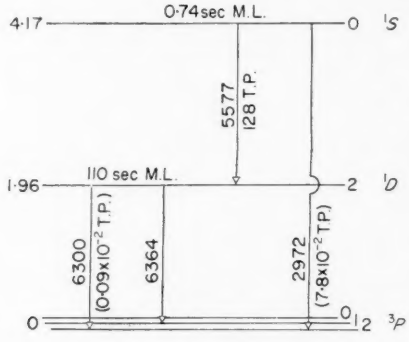
$$z = R(\sec\beta - 1) \tag{4}$$

In the preceding equations, z is the height of intersection of the sun's rays and the line of sight, R is the radius of the earth, β is the solar depression at the point of intersection, θ is the angle of emergence of the radiation being observed, ψ is the zenith distance of observation, α is the solar depression at the observer, and ΔA is the difference in azimuth between the point of observation and the sun. These four equations involve four unknowns and a closed solution is possible. However, the equations are transcendental and in practice the solution is obtained by an iterative process. An initial value of z is chosen (in this case 150 km) and inserted in equation 1 and a resultant value of z calculated; this value is then inserted in equation 1 and the process repeated until successive values agree to within 1 km.

The above calculations involve the assumption that the earth is screening the sun's rays from the upper atmosphere. The direct excitation of λ5577, however,

takes photons of a wave-length <2972 Å as demonstrated in Fig. 4. This wave-length is strongly absorbed by ozone. It is necessary then to calculate the screening height of the ozone. For this purpose a model ozone distribution was chosen which fits reasonably well with rocket data for the upper atmosphere.

An ozone distribution of the form $N(O_3) = N_0(O_3)e^{-\lambda h}$ was assumed where h is measured from the earth's surface (JOHNSON *et al.*, 1954). To obtain the screening height a value of $7{\cdot}5$ cm^{-1} at s.t.p. was assumed for the absorption coefficient of ozone at 2972 Å (FABRY, 1950). The assumption of an exponential distribution is



Partial energy level diagram for [OI]

Fig. 4. This figure shows a partial energy level diagram for atomic oxygen. The mean life is abbreviated as M.L. and the transition probability as T.P.

valid for this problem as only the upper portions of the ozone layer are of interest for this problem. The total absorption was then calculated as a function of the solar depression. The result of this calculation indicates a screening height of 35 km. The height calculated as described above is therefore modified by adding a term $\Delta z = \bar{O}_3 \sec \beta/(1 + \tan \theta \tan \beta \cos \Delta A)$ where $\bar{O}_3$ is the screening height due to ozone.

Refraction in this case gives an altitude correction of less than 1 km and is therefore not included.

The data for both the morning and evening twilights are shown plotted against height in Fig. 5. Intensities at eight azimuths at 75° zenith distance are included in the evening curve and five in the morning. The remainder of the morning readings were not considered valid because of light clouds in those parts of the sky.

The photometer has an acceptance angle of 4°. This implies that the radiation received by the photometer when observing at 75° zenith distance comes from a volume of the sky which is about 30 km in vertical extent, assuming a 100 km height. The height discrimination of this technique is therefore about 25 km. Accordingly, a smooth curve is drawn through the data and values read off in 25 km increments. The differences of these data are presented in Table 2.

INTERPRETATION

The data of Table 2 are plotted in Fig. 6. A least squares fit of the data yields the line drawn in. The scale height for the emission as determined from the curve is

VOL. 17 1959/6

OL.
17
9/60

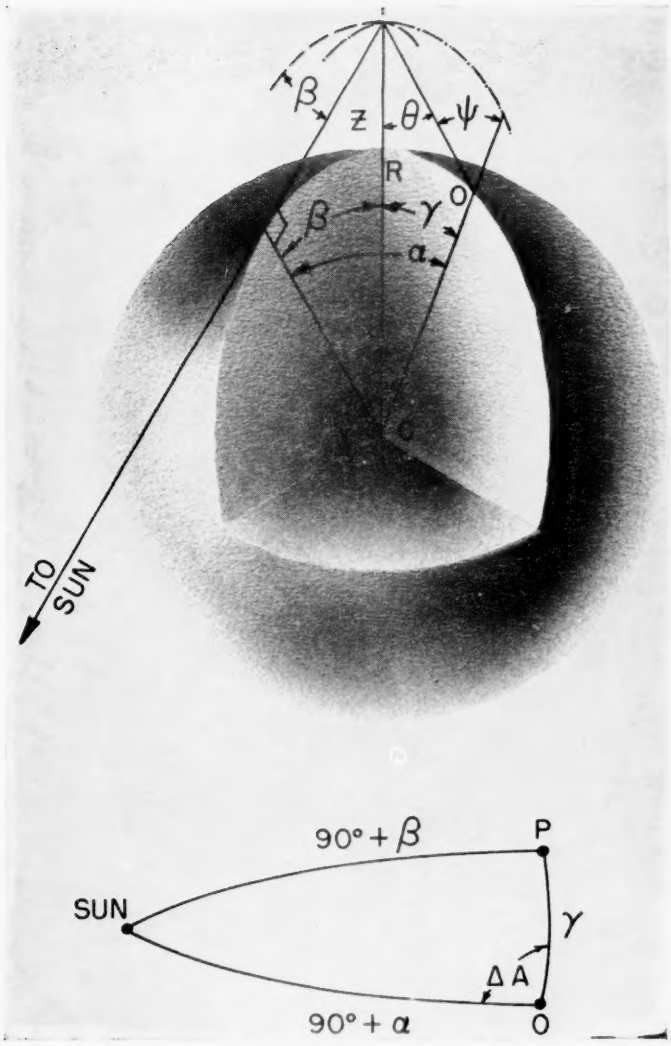
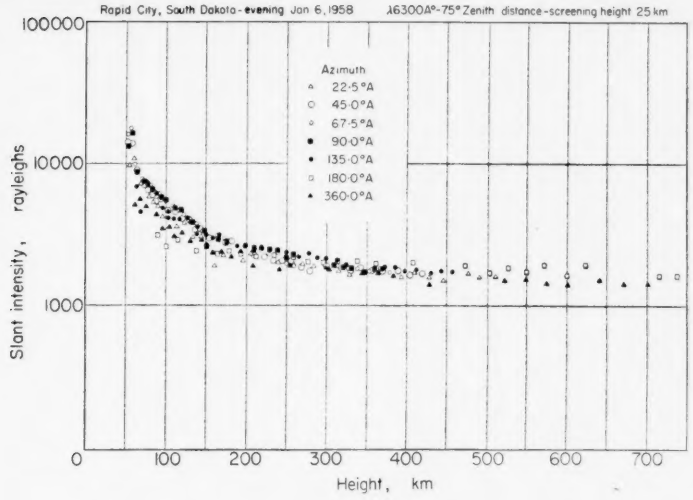
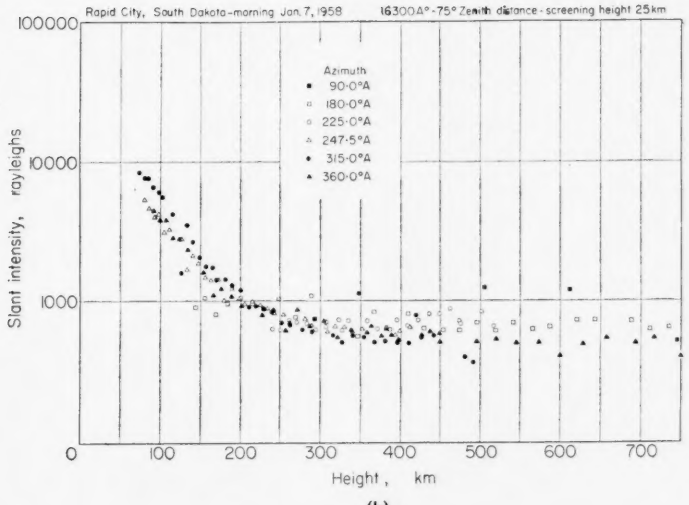


Fig. 3. This drawing demonstrates the geometry needed for the calculation of the intersection heights. *p* Is the point of intersection of the shadow line and the line of sight. The position of the observer is designated as *o* and the centre of the earth as *c*.

VOL.
17
1959/6



(a)



(b)

Fig. 5. This figure shows the log of the slant intensity plotted against height for the λ5577 line. Note that a height of about 100 km is indicated.

28·5 km. Although the distribution of atomic oxygen is not logarithmic in this region, an approximate value for the scale height of atomic oxygen is 12 km. It is not possible to interpret the scale height for the emission as giving the scale height of atomic oxygen however, since deactivation may enter into the reaction rather strongly at lower altitudes which would tend to yield large values for scale height.

It is instructive to make some upper limit calculations concerning the expected emission of λ5577 due to direct excitation by the sunlight. In order to do this, it will be assumed that the sunlight striking the observed layer is unattenuated by its previous passage through the atmosphere and that deactivation is not an important factor.

Table 2

Intersection height	Slant intensity (Rayleighs)		Incremental intensity	
	Evening	Morning	Evening	Morning
250	700	700		
225	700	700	0	0
200	700	700	0	0
175	730	780	30	80
150	860	910	130	130
125	1180	1120	320	210
100	1950	1610	770	490
75	4400	3200	2450	1590

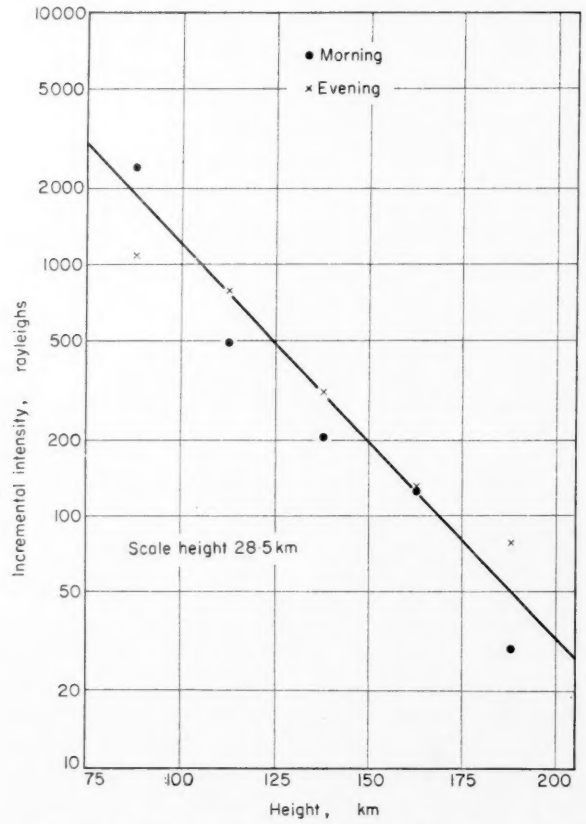


Fig. 6. This figure shows the log intensity per 25 km layer plotted against height in the atmosphere.

Then the number of atoms per cubic centimetre per second excited to the ^1S state will be given by

$$N^*(\mathrm{O}) = N(\mathrm{O})\pi F(\nu)\alpha(\nu)$$

where $F(\nu)$ is the solar flux per square centimetre normal to the beam. $N(\mathrm{O})$ is the number of O atoms per cubic centimetre capable of absorbing frequency ν. $\alpha(\nu)$ Is the absorption cross-section. Of these, a fraction $A_{sd}/(A_{sd} + A_{sp}) \simeq 1$ will radiate a λ5577 quantum if deactivation is neglected. The As are transition probabilities for the transitions indicated. Integration over the λ5577 line, assuming only Doppler broadening, yields for the intensity per centimetre-thick layer:

$$I = F(\nu)(\pi e^2/mc)f_{sd}N(\mathrm{O})$$

f_{sd} is 13×10^{12} and $F(\nu)$ about 25 photons.cm^{-2}.sec^{-1} (Allen, 1955). $N(\mathrm{O})$ at 125 km is taken as 180×10^{10} atoms.cm^{-3} (Seaton, 1958). Under these conditions, after converting to Rayleighs per 25 km thick layer observed at a zenith distance of 75°, one obtains a value of about 100 Rayleighs as the expected intensity. The measured intensity is about 500 Rayleighs. It would seem then that direct excitation by sunlight can, at best, explain only a fraction of the enhancement, and another mechanism must be looked for.

A clue as to the mechanism may be found in the scale height for the emission. If the mechanism involves free electrons, one would expect to measure a scale height quite different than that for oxygen alone. The difference will depend on the scale height of the electrons and the mechanism involved. The scale height measured is slightly more than twice the scale height for oxygen atoms.

OL.
17
9/60

The 6300 Data

The λ6300 data were analysed in the same manner as the λ5577 data except that a screening height of 25 km was used. The data show a marked dissimilarity between the twilight and dawn (Fig. 7). Consequently, one does not expect the explanation of the enhancement to be due to resonance scattering. Chamberlain (1955) has presented a theory for the λ6300 emission. He postulates that the principal contribution comes from change transfer from O^+ to O_2 with a subsequent recombination of O_2^+ and an electron resulting in two excited oxygen atoms as follows:

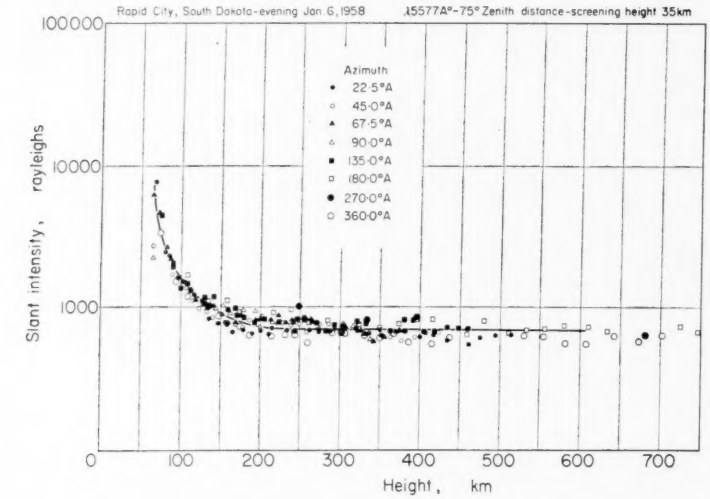
$$O_2 + O^+ \rightarrow O_2^+ + O$$

$$O_2^+ + e \rightarrow O^* + O^*$$

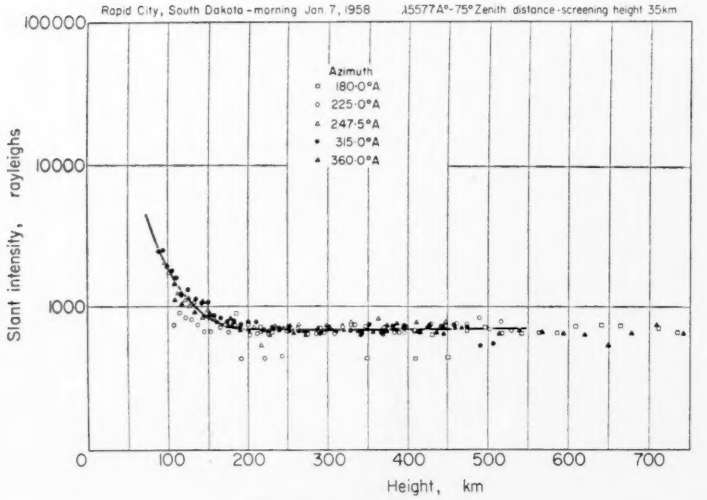
Fig. 8 compares the results of Chamberlain with the experimental results. At first glance there seems to be a rather poor comparison between theory and observation. The general variation with time, however, is similar for long times after sunset. Also, many of the rate coefficients involved in the calculation are not well known. It should also be pointed out that the average intensity during this night is somewhat higher than is usually present.

Acknowledgments—Especial credit for this work must go to Mr J. E. CRUZ who operated the Airglow IGY observing site at Rapid City, South Dakota, for the South Dakota School of Mines and Technology. He not only was the observer on the night analysed for this paper, but was the first to point out the enhancement of $\lambda 5577$ line in the twilight.

Thanks are also due to Dr. F. E. ROACH, Mr. E. TANDBERG-HANSSEN and Miss P. JAMNICK who have aided me with valuable discussions and encouragement.



(a)



(b)

Fig. 7. A plot of the $\lambda 6300$ radiation slant intensity against height is shown here. Note that the evening intensity decays much more slowly than does the morning, indicating a long time constant for the excitation process.

VOL.
17
1959/6

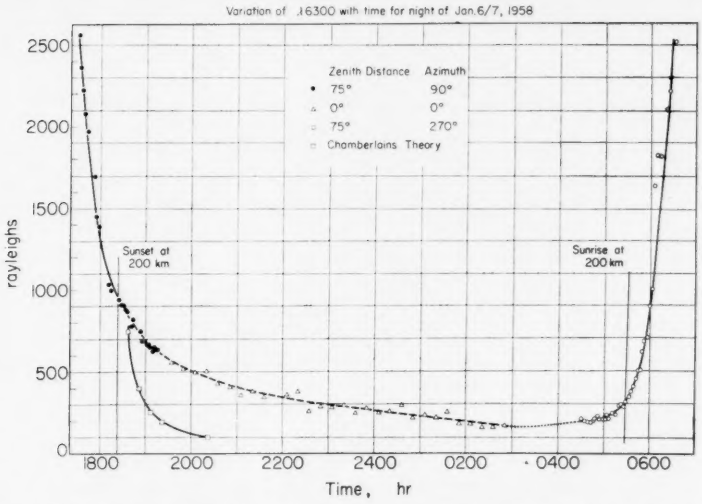


Fig. 8. This plot is a composite of three sets of readings put together to give the longest possible history for the λ6300 emission. The early part of the curve is derived from observations in the west at 75° zenith distance, corrected to the zenith. The central portion of the curve is zenith data and the later portion is derived from observations in the East at 75° zenith distance, corrected to the zenith. The general intensity during this night is somewhat higher than is common. The theoretical curve predicted by CHAMBERLAIN is drawn in.

REFERENCES

ALLEN C. W.	1955	*Astrophysical Quantities.* Athlone Press, London.
ASHBURN O.	1954	*J. Atmosph. Terr. Phys.* **5,** 83.
BLAMONT J. E. and KASTLER A.	1951	*Ann Geophys.* **7,** 73.
DUFAY J. and DUFAY M.	1948	*C.R. Acad. Sci., Paris* **220,** 1209.
DUNN R. B. and MANRING E. R.	1955	*J. Opt. Soc. Amer.* **45,** 899.
FABRY C.	1950	*Editions on Centre de la Recherche Scientifique* p. 106.
JOHNSON F. S., PURCELL J. D. and TOWSEY R.	1954	*Rocket Exploration of the Upper Atmosphere* (Edited by BOYD R. L. F. and SEATON M. J.). Interscience, New York: Pergamon Press, London.
KOOMEN M. J., PACKER D. M. and TOWSEY R.	1955	*The Airglow and the Aurorae* (Edited by ARMSTRONG E. B. and DALGARNO A.) p. 355. Pergamon Press, London.
MANRING E. R. and PETIT H. B.	1955	*The Airglow and the Aurorae* (Edited by ARMSTRONG E. B. and DALGARNO A.) p. 29. Pergamon Press, London.
ROACH F. E.	1958	*The Intercalibration of Airglow Photometers.* NBS Report 5591.
SEATON M. J.	1958	*Astrophys. J.* **127,** 67.

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 286 to 294. Pergamon Press Ltd. Printed in Northern Ireland

Fluctuations in the geomagnetic horizontal field near the magnetic equator

A. Onwumechilli
Physics Department, University College, Ibadan, Nigeria

(*Received* 6 *July* 1959)

Abstract—Six-hundred fluctuations in the geomagnetic horizontal component (H) and vertical component (Z) have been measured from the magnetograms taken at Ibadan in April and May 1958. An analysis of these has shown that: (a) the frequency of occurrence varies with local time in the same way as S_q; (b) the amplitude of fluctuations increases in a general way with S_q; (c) the amplitude of fluctuations in both H and Z tend to increase linearly with period in duration (T) and there is some evidence for a change of slope at about $T = 33$ min with the fluctuations of longer duration increasing more slowly with T than those of shorter duration; (d) fluctuations under quiet conditions occur mostly during day-time and are more frequent when S_q is high; (e) fluctuations during the night are of long duration and mostly occur during disturbance ($K_p > 3+$). It has been shown that fluctuations under quiet conditions have the characteristics of solar-flare effects and it is therefore suggested that both arise from similar causes.

1. Introduction

Chapman and Bartels (1940) observed that even on quiet days there is considerable fluctuation in the geomagnetic horizontal component at Huancayo near the magnetic equator. Onwumechilli (1959a,c) initiated the study of these fluctuations and later widened it (Onwumechilli, 1959b) to cover all fluctuations so that whatever differences that may exist between types will be appreciated. In this paper the study is continued to cover the diurnal characteristics of the fluctuations and to show in what respects the fluctuations may be compared with solar-flare effects, for it has been suggested by Onwumechilli in the above-mentioned papers that they are similar phenomena.

The data comprise 600 fluctuations measured on the magnetograms of Ibadan, Nigeria, for April and May 1958. For each fluctuation the following parameters were measured—the amplitude of the fluctuation (m) being the maximum departure from the current level of the diurnal variation; the period in duration (T) being the time interval between the departure from and the return to the trend of the diurnal variation; the instantaneous value of the diurnal variation (M_t) being the mean of the values of H at the beginning and the end of the fluctuation minus the mean of the hourly values of H at the preceding and succeeding midnights; and the daily range (M) measured in the usual way. Also each fluctuation is assigned to the hour nearest its maximum departure and the K_p index at the time of occurrence is recorded. All easily measurable fluctuations were included, the limit being $|m| \geqslant 1$ mm (i.e. 3·74 gammas) and $T \geqslant 2$ mm (i.e. 8 min). The data have been described in greater detail in a previous paper (Onwumechilli, 1959b).

2. The Frequency of Occurrence

In Fig. 1 the frequency of occurrence is plotted against local time for various lower limits of the amplitude of the fluctuations. It shows clearly that indentations of all amplitudes are more numerous on a magnetogram during the day than the

night. (The dip in Fig. 1 at 0900 hours may be excused. It is likely due to loss of records during the routine changing of films at that hour). The shape of the curve is that of the diurnal variation of H-minima around sunrise and sunset and maximum around noon. About half the indentations in each hour have amplitudes between 3 and 11 gammas.

To study the effect of disturbance on the fluctuations, two special groups were selected according as they occurred during periods with $K_p \leqslant 3_0$ (quiet) or

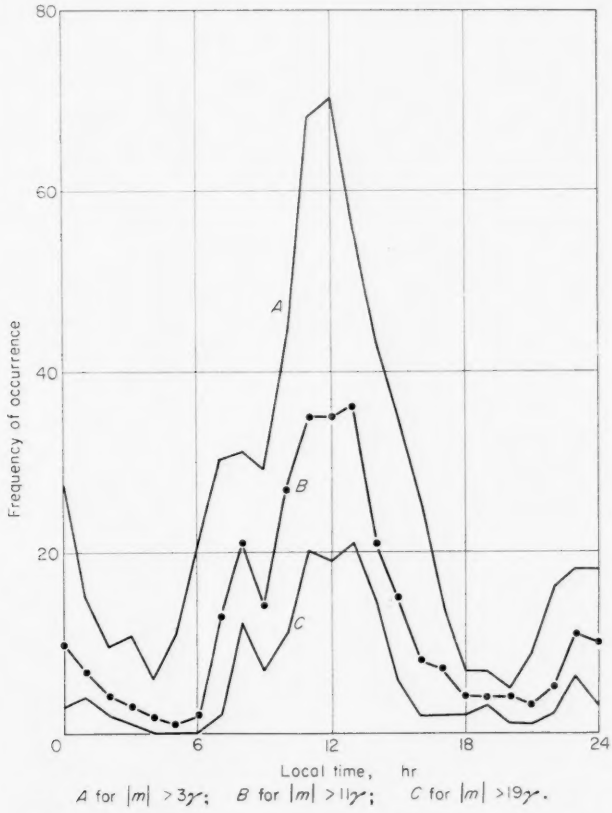


Fig. 1. Frequency of occurrence of fluctuations in H-variations.

$K_p \geqslant 4+$ (disturbed). The lower limit for these groups is $|m| \geqslant 11$ gammas (3 mm) since it is not likely that any indentation of amplitude 3 mm on the magnetogram would be neglected during the measurements.

The division affected the frequency of occurrence remarkably. None of the fifty fluctuations under quiet condition occurred in the night. They all came between 0700 and 1500 hours LT. The shape of the occurrence frequency curve is similar to Fig. 1 and the statements would still hold even if fluctuations of $K_p = 3+$ were included (bringing the total to 80) see Fig. 2. On the other hand the 131 in the disturbed group occurred during both day and night.

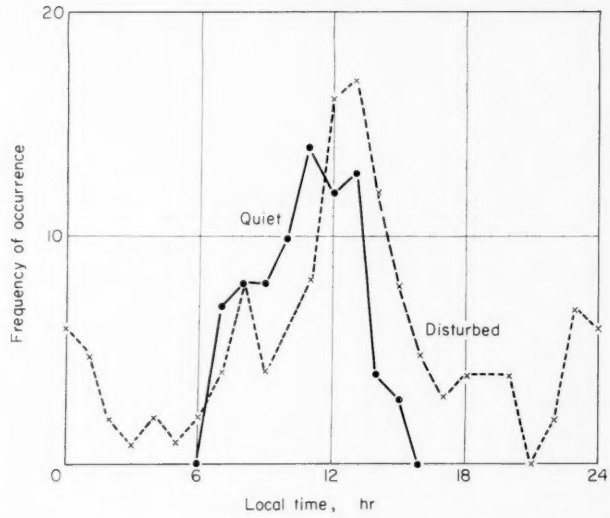


Fig. 2. Frequency of occurrence of fluctuations in H during quiet and disturbed conditions.

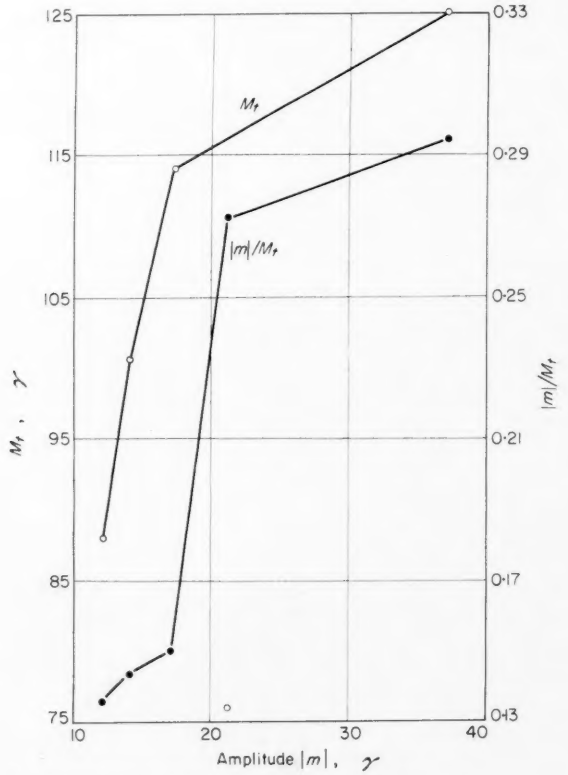


Fig. 3. Dependence of M_t and $|m|/M_t$ on $|m|$ under quiet conditions.

3. The Amplitude of the Fluctuations

The quiet group is arranged into subdivisions (of ten each) in ascending order of magnitude of $|m|$. Then $|m|$, M_t and T, respectively, are averaged for each group. Fig. 3 demonstrates that the mean of $|m|$ increases in a general way with the means of M_t and $|m|/M_t$, respectively. The increase of $|m|$ with M_t may be expected from Fig. 1 and an independent confirmation of the result is obtained from the finding that the average of $|m|$ for each hour is highest around noon and varies with local time in a manner similar to Fig. 2.

The eighty-four fluctuations occurring in day-time (0800 to 1600 hours) under disturbed condition are also averaged in groups of ten in ascending order of magnitude of $|m|$. The mean of $|m|$ increases with duration T but no clear evidence of a similar increase with M_t is obtained. This is not enough for stressing the lack of evidence because M_t varies widely during disturbance and these small samples of ten may not be expected to reveal any weak tendency, but the case for lack of evidence is strengthened by the failure of the hourly means and medians of $|m|$ to show a definite trend in their connexions with time of day.

It was also found that $|m|$ shows no evidence of dependence on K_p.

On treating the thirty-four fluctuations under disturbance during the night (1900 to 0500 hours) like the others, it was found that $|m|$ increases with duration T. On the whole the fluctuations during the night are of long periods of duration (average $T = 70$ min).

OL.
17
9/60

4. Dependence of Amplitude on Duration

The averages of T for the groups of ten in Section 3 show that $|m|$ tends to increase with T during both quiet and disturbed conditions. This requires more detailed investigation because it might give an insight into the unexplained finding that the ratio of the amplitude of a fluctuation in H to the amplitude of the corresponding fluctuation in Z decreases with increase in T. The linear relation between T and this ratio also showed a change of slope at about $T = 35$ min (Onwumechilli, 1959b) but this was not commented upon then. Since a similar change of slope has been reported by Jacobs (1959) from an investigation of a similar phenomenon, it has been taken into account in the following study.

Some 342 fluctuations of all magnitudes, occurring under varying conditions of disturbance between the hours of 0800 and 1400, were pooled together for the investigation of the dependence of $|m|$ on T. The data was divided into two groups—234 with duration $T < 30$ min in group A and 79 with $T \geqslant 35$ min in group B. The regression line of $|m|$ on T,

$$|m| = a + b(T - \bar{T}) \tag{1}$$

was calculated for each group. $\bar{T}$ is the average of T for the group. The results are given in Table 1.

To display the regression, the data are arranged in ascending order of magnitude of T and averaged in groups of about 40 for $T < 30$ min and groups of about 20 for $T \geqslant 35$ min. These averages are the points plotted in Fig. 4. The four regression lines are also drawn in the same figure.

Table 1. Showing the coefficients of the regression of $|m|$ on T

	a	b	T
Fluctuation in H group A	13·25	0·5019 ± 0·0648	17·73
Fluctuation in H group B	28·61	0·3805 ± 0·0506	53·55
Fluctuation in Z group A	6·17	0·2635 ± 0·0280	17·73
Fluctuation in Z group B	16·28	0·1976 ± 0·0306	53·55

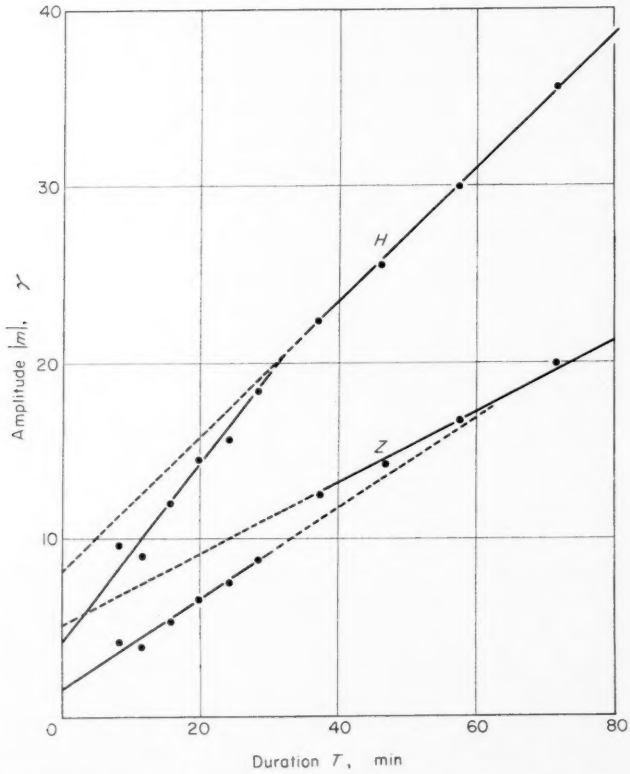


Fig. 4. Showing dependence of the amplitude on the duration of fluctuations in H and Z.

It is clear from Fig. 4 that on the average the amplitude of fluctuations in both H and Z tend to increase linearly with duration T. The rate of increase with T is smaller for fluctuations of long duration ($\geqslant$35 min) than for those of short periods.

5. Comparison of the Indentations with Crochets

One difference between the crochets and indentations is that, while the former are always increases in H, the latter may be increases or decreases. For further

VOL. 17 1959/6

comparison let us consider first only those fluctuations occurring under quiet condition.

The comparison has to be based on what the earlier workers (McNish, 1937; Nagata, 1952; Forbush, 1956) regarded as the characteristics of solar flare effects or crochets. Like crochets the fluctuations are most numerous around noon but the most striking similarity is that both only occur during daylight hours. At Huancayo, Nagata (1952) found crochets between about 0500 and 1700 hours LT and Forbush (1956) between 0700 and 1500 LT and here all the 80 fluctuations come between 0700 and 1500 hours LT. The variation of the

Table 2. Comparing $|m|/M_t$ for crochets and fluctuations

Type	Count	Mean $\|m\|$	Mean $\|m\|/M_t$	Mean T (min)
Crochet				
(Huancayo)	74	25·8	0·283	26·3*
(Hakioka)	62	7·2	0·296	24·8*
(Ibadan)	22	28·1	0·352	28·9
Fluctuation				
(Ibadan)	50	18·1	0·212	30·8
(<Median)	25	12·9	0·103	24·4
(>Median)	25	24·4	0·294	38·9

* Sum of Nagata's T_1 and T_2.

frequency of occurrence with time of day (Fig. 2) is also very similar to that found for crochets by Nagata (1952).

Forbush (1956) found that crochet amplitude (m) tends to increase in a general way with S_q magnitude (M_t) and that on the average they both vary in the same way with LT. This is the same as the finding here for fluctuations under quiet condition. Since $|m|$ increases with M_t (Fig. 3), it follows that, as M_t (or S_q) increases towards local noon, $|m|$ would increase. This has also been verified directly—the hourly means of $|m|$ is highest around local noon like S_q.

To bring the comparison of fluctuations and crochets nearer home, the obvious crochets at Ibadan from November 1955 to April 1959 have been measured. The number (22) is too small for most statistical analysis but is, however, examined for information additional to that obtainable from other workers on crochets. These 22 crochets occurred between 0900 and 1700 hours LT and all during periods with K_p index not exceeding 3+.

Nagata (1952) studied the duration T and the ratio m/M_t as characteristics of crochets. The most probable values of these (as given by the mean of their logarithms) together with other relevant data are given in Table 2. The data for Huancayo and Kakioka are from Nagata. m/M_t for crochets at Ibadan is higher than the same ratio for Huancayo and Kakioka. The most likely reason is the higher value of $|m|$ at Ibadan for it has been shown (Fig. 3) that $|m|/M_t$ increases with $|m|$. The same ratio for fluctuations at Ibadan is slightly less than those of Huancayo and Kakioka probably because $|m|$ for the fluctuations is small. The

argument is exemplified by taking only those fluctuations whose amplitudes are greater than the median value. For this group the average $|m|$ is almost the same as at Huancayo and the ratio $|m|/M_t$ also becomes almost the same as at Huancayo. Moreover, the histogram of $|m|$ is found to be similar to that of Nagata (1952). It must therefore be concluded that as far as the so-called characteristics of crochets can be relied on, the fluctuations under quiet conditions are like crochets or geomagnetic solar-flare effects.

For completeness we may note that taking all the fluctuations together, the most probable duration T agrees very well with that of crochets. The duration for the fluctuations greater than the median is, however, higher than the duration for crochets.

On the other hand, when the fluctuations under disturbed conditions are compared with crochets we get both disagreements and similarities. Unlike crochets the fluctuations occur at all hours of the day. Although they are more numerous during day-time than night-time, there is little evidence that the amplitude of the fluctuation depends on local time or on the strength of S_q. For the hours of 0800 to 1600 LT the mean $|m|/M_t$ is 0·276 and the mean duration is 30 min. These agree well with the same values from crochets but the agreement loses its significance if the fluctuations are not in other ways similar to crochets.

In conclusion it may be said that fluctuations under quiet conditions have the same characteristics as crochets. The evidence in this paper does not justify the same statement concerning fluctuations under disturbed conditions.

VOL.
17
1959/6

6. Discussions of Results

Geomagnetic variations are caused by changes in the strength and direction of ionospheric currents. Such variations are seen in a better perspective when they are studied from two or more geomagnetic field components. Therefore, in discussing the results presented here, the relationship between H and Z variations at Ibadan already reported by Onwumechilli (1959b) is kept fully in view, for example, since it was shown that the ratio of the amplitudes of indentations in H to the amplitudes of the corresponding indentations in Z is constant between the hours of 0800 and 1600 LT it is assumed here that the results obtained for fluctuations in H between those hours will in general apply to Z fluctuations.

In discussing geomagnetic variations under the purview of the "dynamo" theory, both the ionization and the wind system at the level of the overhead current are very important, and such a discussion is complicated by the fact that the magnitude of an effect caused by one equally depends on the other. It is therefore simpler if discussions of magnitude and cause are separated as much as possible.

The greater occurrence of fluctuations during day-time must be due to the fact that the magnitude of the effect depends on the S_q current system for, since the amplitude of the fluctuations increases with S_q (i.e. M_t), weaker causes have greater chances of their effects being observed around noon when the S_q current system is strongest. During the night S_q and its current system are low and only strong causes have any chance of making their effects observable. This conclusion from occurrence frequency is supported by the fact that fluctuations during the night

are of long duration—the longer the duration the larger the amplitude (see Fig. 4). That hourly averages of amplitudes during quiet conditions are greatest around noon is a further confirmation of the same idea.

The strength of the current system depends as much on the ionization density as on the strength of the prevailing wind but its direction depends more on the direction of the wind. The relative effects of the current on H and Z depend on its direction. It was therefore concluded that since the ratio of the effect on H to that on Z is constant between 0800 and 1400, the direction of the prevailing wind also remains fairly constant at the same time (ONWUMECHILLI, 1959b). Now we know further that the amplitudes of both H and Z fluctuations tend to increase linearly with duration. The relative rates of increase as given by the ratio of the gradient of the regression line for H to the gradient of the regression line for Z is 1·91 for $T < 30$ min and 1·92 for $T \geqslant 35$ min. It is thus clear that the ratio (r) of the amplitude of a fluctuation in H to the amplitude of the corresponding fluctuation in Z decreases with duration T because $(\Delta m)_H/(\Delta m)_Z < r$ where $(\Delta m)_H$ and $(\Delta m)_Z$ are increases in the amplitudes of H and Z fluctuations, respectively. To keep r constant $(\Delta m)_H/(\Delta m)_Z$ must equal r. In support of this explanation we note that r varies little after decreasing to the value 1·91 at $T = 35$ min (ONWUMECHILLI, 1959b).

Some reasons may be suggested for $(\Delta m)_H/(\Delta m)_Z$ being less than r. About 50 per cent of the variation in H at Ibadan is due to the effect of the electrojet while as much as about 80 per cent of the variation in Z is due to the same cause (ONWUMECHILLI, 1959c). If therefore the electrojet fluctuates disproportionately more than the normal S_q current system, the result would be a proportionately larger fluctuation in Z than in H. An equally plausible suggestion is that as the amplitude of fluctuations is caused to increase, the direction of the prevailing wind may also be changing. A change from the optimum horizontal wind direction for H would tend to reduce the amplitude of H fluctuations without affecting the Z amplitudes so much. The doubt here is that m is as often positive as it is negative so it is difficult to see why such changes in wind direction should always affect H adversely. It is hoped that the investigation of the spatial distribution of the fluctuations will shed some light on this problem.

The increase of $|m|$ with T does not appear to be of any intrinsic interest since a cause lasting longer may be expected to be stronger and to produce stronger effect. Of interest, however, is the change of slope in the linear variation of $|m|$ with T. It occurs at about $T = 33$ min. JACOBS (1959), in the investigation of geomagnetic micropulsations found a similar change of slope at $T = 2$ min and he attributed it to the existence of two different kinds of magnetohydrodynamic waves. A similar cause could be considered here but it is felt that the change of slope should first be proved significant before further consideration. Even though the difference between the gradients of the regression lines for groups A and B is about twice the standard error (Table 1), the more rigorous "Student's" test (FISHER, 1954, p. 140) shows that the change of slope is not statistically significant.

It has been shown in Section 5 that fluctuations under quiet conditions have the same characteristics as solar-flare effects. The natural suggestion is that both could arise from the same cause. This would mean that fluctuations under quiet

conditions are caused by fluctuations in the quantity and quality of ionizing wave radiations from the sun. The magnitude of the effect would ultimately depend on the ionospheric conductivity and the prevailing wind. Because the former is particularly large near the magnetic equator, the fluctuations would be enhanced in that region.

It is more difficult to say how fluctuations under disturbed conditions arise. If they are due to solar radiations, these must be corpuscular since they occur at night. It is remarkable that the ratio of the amplitude of a fluctuation in H to the corresponding fluctuation in Z is independent of disturbance (ONWUMECHILLI, 1959b). It means that even if the currents producing fluctuations on a disturbed day flow at a height different from that for the fluctuations on a quiet day, they must at least have identical directions. More is in fact expected for, if their heights are different, the enhanced electrojet effect on Z at Ibadan should have made the ratio for quiet conditions smaller than that for disturbed conditions. This will be further discussed in another paper on the spatial distribution of the fluctuations.

REFERENCES

CHAPMAN S. and BARTELS J.	1940	*Geomagnetism* p. 242. Oxford University Press.
FISHER R. A.	1954	*Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.
FORBUSH S. E.	1956	*J. Geophys. Res.* **61,** 93.
JACOBS J. A.	1959	*Nature, Lond.* **183,** 381.
MCNISH A. G.	1937	*Terr. Magn. Atmos. Elect.* **42,** 109.
NAGATA T.	1952	*J. Geophys. Res.* **57,** 1.
ONWUMECHILLI C. A.	1959a	*J. Atmosph. Terr. Phys.* **13,** 222.
ONWUMECHILLI C. A.	1959b	*J. Atmosph. Terr. Phys.* **16,** 274.
ONWUMECHILLI C. A.	1959c	*J. Atmosph. Terr. Phys.* **13,** 235.

VOL. 17 1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 295 to 301. Pergamon Press Ltd. Printed in Northern Ireland

Resonance scattering by atmospheric sodium—VIII

An improved method of deducing the vertical distribution*

D. M. HUNTEN
University of Saskatchewan, Saskatoon

(*Received* 7 *July* 1959)

Abstract—It is shown that the derivative of sodium twilight brightness with shadow height may be regarded as a smeared version of the vertical distribution of sodium atoms. The distribution may be recovered by use of the chord construction of BRACEWELL. The method is tested on a model distribution and found satisfactory; it is easier to use and gives more detail than the matrix method previously proposed. When the sodium abundance is large, the derivative curve must be corrected for resonance absorption before it can be used. A simple method of doing this is described and the required numbers tabulated.

1. INTRODUCTION

THE problem of deducing the vertical distribution of atmospheric sodium from twilight measurements of the *D*-lines is one of long standing. It was discussed in some detail in Paper VII of this series (RUNDLE *et al.*, 1959) where a matrix method of solving the integral equation was developed and tested. The results were useful but not wholly satisfactory. The present paper points out that the integral equation is analogous to the one which is found in smoothing problems in such fields as radio astronomy (BRACEWELL and ROBERTS, 1954) or spectral scanning with inadequate resolution (HARDY and YOUNG, 1949). A recent paper by DIXON and AITKEN (1958) has shown that the matrix method is apparently not suited to such problems, and our experience discussed in Paper VII seems to confirm this.

The new viewpoint is to regard the transmission derivative $\dot{T}$ as sweeping over the sodium distribution n to generate the brightness derivative $\dot{B}$ according to equation 1; thus $\dot{B}$ is regarded as a smeared version of n and the problem is to remove the smearing by a suitable restoring process. The references just mentioned are only a few of a large literature on that subject, and known methods may immediately be applied to our problem. In particular, we will use the elegant graphical method of BRACEWELL (1955) which is very easy to apply. BRACEWELL and ROBERTS (1954) have shown that the corrected result is not always the same as the original because the integral equation may not have a unique solution. We have encountered some difficulties of this sort, and have been able to alleviate them by subtracting from the observed $\dot{B}$ a standard curve which is the $\dot{B}$ for an assumed distribution; restoration then need be carried out only on the difference. The result is, of course, biased towards the assumed one, but this may be preferable to the other difficulties.

* The research reported in this paper was supported by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)–1831.

When the sodium abundance is large, resonance absorption may have an appreciable effect on the observed $\dot{B}$; this effect was discussed in Paper VII, but the only result quoted was that it could be neglected for the lower abundances found in practice. A simple and compact method of presenting the results has now been devised, and they are given in graphical and tabular form. With their aid the observations may be corrected to the shape which they would have in the absence of resonance absorption.

2. Chord Treatment of the Integral Equation

The equation to be treated is (2) of Paper VII; omitting a factor as in that paper this is

$$\dot{B}_s(z_1) = \int_{z_1 - 25}^{\infty} \dot{T}(z - z_1) n(z)\, dz \qquad (1)$$

The dot represents differentiation with respect to z_1, the geometrical shadow height. As already mentioned, this equation can be regarded as describing the effect of scanning a function $n(z)$ with an instrumental profile $\dot{T}$ to produce an observed function $\dot{B}_s$; BRACEWELL's equation (1) (1955) is essentially the same. It is convenient to have $\int \dot{T}\, dz = 1$; here this is assured since $T(-25) = 0$ and $T(\infty) = 1$. In most smoothing problems the first moment of $\dot{T}$ vanishes; here this is not so unless the origin is shifted up by an amount which we call $\bar{x}$, where

$$\bar{x} = \int_{z_1 - 25}^{\infty} x \dot{T}(x)\, dx$$

and $x = z - z_1$. We will also need to know the second moment σ^2, $\bar{x}$ and σ^2 were evaluated numerically for the "spring" and "autumn" transmission functions used in Paper VII (the variation is caused by the seasonal changes in ozone). The results were:

spring:	$\bar{x} = 27{\cdot}0$ km,	$2\sigma = 19{\cdot}4$ km
autumn:	$26{\cdot}3$ km,	$21{\cdot}2$ km

Treatment of the observations according to the BRACEWELL method is carried out as in Fig. 1, which shows the results of a test calculation. Plot $\dot{B}_s(z_1)$ from observation against $(z_1 + \bar{x})$; z_1 follows from the solar depression. The result is to be considered as a broadened version of the sodium distribution $n(z)$. The sharpening procedure does not appreciably change the position of the peak; this can therefore be seen immediately. A chord is drawn across the curve spanning a height-range $2a = 2\sigma$; the correction to the curve at the middle of the chord is equal to the distance from the chord to the curve. Since the shift $\bar{x}$ and the span $2a$ vary with the ozone abundance, they must be chosen to suit the season. The variations are small and the exact choice makes little difference to the final result; thus, the method is happily insensitive to such variations and to other errors in the calculated $\dot{T}$. There is a slight asymmetry in $\dot{T}$ which is not taken into account; the effect of this can be seen in Fig 1 and is fortunately slight.

The model distribution shown in Fig. 1 was used to test the method. The models previously used contained discontinuities in slope which seem to upset

VOL.
17
1959/6

the results; therefore the smooth curve $n(z) = \text{sech}\ [(z - 85)/7{\cdot}5]$ was used (lengths in kilometres). This function goes over into exponentials at heights well above and below the peak, and should be similar to the atmospheric distribution of sodium. $\dot{B}_s(z_1)$ was calculated by electronic computer using the same programme as in Paper VII; it is shown also in Fig. 1 plotted against $(z_1 + 27$ km) and the resemblance to a broadened version of $n(z)$ may be seen. The curve corrected by the chord method agrees fairly well with the original; it falls somewhat short at the peak, but the relative error is not large. The undershoots about 25 km above

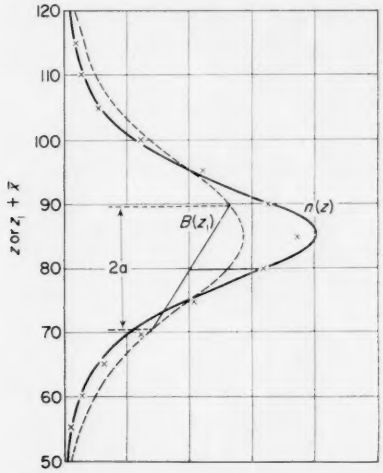


Fig. 1. Test of the chord method. From the assumed distribution $n(z)$ the corresponding brightness derivative $\dot{B}(z_1)$ was calculated. The chord method then gave the restored distribution shown by the crosses.

and below the peak are more annoying. They do not seem to be due to omission of higher-order terms of the correction series, since inclusion of the fourth-order term as suggested by BRACEWELL (1955) makes them worse. It seems rather that the slightly oscillating solution is mathematically just as good as the original distribution; this is possible because equations like (1) may not have a unique solution (BRACEWELL and ROBERTS, 1954). These authors give several examples of similar situations.

The difficulty seems to be caused by the presence of a fairly sharp peak in the distribution; it may be alleviated by subtracting an assumed solution so that the remainder is better behaved. A suitable solution to subtract is the one in Fig. 1; let this be called $n_0(z)$ and the corresponding brightness derivative $\dot{B}_0(z)$. Then the actual distribution $n = n_0 + \Delta n$; substitute this in (1):

$$\dot{B}_s(z_1) = \int \dot{T} n_0 \, dz + \int \dot{T} \Delta n \, dz = \dot{B}_0(z_1) + \int \dot{T} \Delta n \, dz.$$
$$\Delta \dot{B}(z_1) \equiv \dot{B}_s - \dot{B}_0 = \int \dot{T} \Delta n \, dz. \qquad (2)$$

Equation (2) is of the same form as (1) and may be solved for Δn in the same way that (1) is solved for n. $\dot{B}_0$ may be moved up or down as well as changed in

amplitude to get the best agreement with the observed $\dot{B}_s$; then n_0 must be treated in the same way. When Δn has been found by the chord method it is added to this modified n_0 to give the observed distribution. There is a tendency for this to resemble n_0, but this may be less serious than the oscillations produced by direct application of the chord method. Perhaps a combination of (1) for the region near the peak and (2) for the tails would be best.

If the subtraction method is used on the curve of Fig. 1, Δn will be zero and the final result perfect. To obtain some idea of the behaviour of the method, a somewhat distorted version of the distribution was drawn and the corresponding $\dot{B}_s$ computed, using Fig. 1 for n_0; the result was very satisfactory and the oscillations were almost completely eliminated. Comparison of the direct and difference methods on experimentally measured $\dot{B}_s$ curves shows that they give almost indistinguishable results; undershoots are often present but they are often much larger than the ones found in Fig. 1. Apparently the defects in the observations or the assumed value of $2a$ are much larger than the defects in the chord method; thus, there is little use in including the refinement embodied in (2). The undershoots often seem to be due to a tendency to draw a curve through a noisy trace which falls too quickly to the axis; great care must be taken to draw a curve which is suitable for the sharpening procedure when the signal/noise ratio is low.

3. Correction for Resonance Absorption

The work described in this section is a continuation of Section 3 of Paper VII. Some light is lost by resonance absorption in the sodium layer on the sunlit and twilit sides of the earth. $\mathscr{S}(z_1)$ is the factor by which the intensity calculated by simple means must be multiplied to allow for this effect; it is a number usually between 0·5 and 1 for the sodium abundances encountered. It was shown that the effect of $\mathscr{S}$ on the shape of the derived distribution could be taken into account by the following expression:

$$\dot{B}_s = \dot{B} - \dot{\mathscr{S}} B / \mathscr{S} \qquad (3)$$

in which B and $\dot{B}$ are the observed brightness and its derivative, and $\dot{B}_s$ is the corrected derivative to be used in (1).

$\dot{\mathscr{S}}$ is a function of N, the total abundance of sodium, but plots of $\dot{\mathscr{S}}$ against z_1 for various values of N appear to have the same shape, differing only by a factor. This was tested by plotting several of them together after normalizing each to 1 at $z_1 = 45$ km. Even for an abundance as large as 17×10^9 atoms/cm^2, these curves almost superposed; the one for 7×10^9 atoms/cm^2 was taken as representative for further work. The correction term in (3) contains also B as well as $\mathscr{S}$ which is taken constant; it is accurate enough to use a standard shape for $B(z_1)$, such as the one from the model described in Section 2. Fig. 2 shows the shapes of $\dot{\mathscr{S}}$ and the product $\dot{\mathscr{S}} B$; the latter is taken as equal to 1·0 at $z_1 = 43$ km. It is seen that a simple straight line gives a satisfactory representation of $\dot{\mathscr{S}} B$; this, along with the several other approximations in this development, is justified by the fact that we are dealing with a small correction. This straight line goes to zero

VOL. 17 1959/6

at $z_1 = 56$ km and has the value given in Table 1 at 43 km; the method of deriving these numbers is described in the Appendix. The abundance N to be used in Table 1 is found from the brightness at a depression of 6° 30′ according to the theory of Paper IV (CHAMBERLAIN *et al.*, 1958) or the equivalent by Donahue and collaborators.

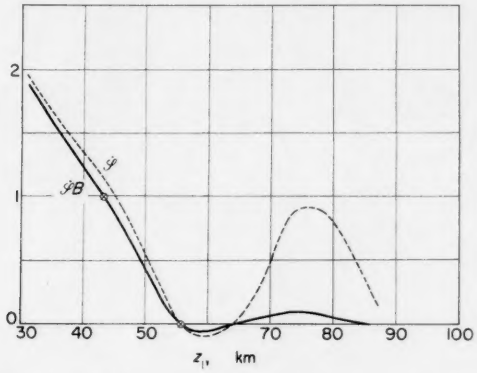


Fig. 2. Shapes of $\dot{\mathscr{I}}(z_1)$ and of $\dot{\mathscr{I}}B$. The two points define the straight line which is used as an approximation to the latter.

Though it has not been mentioned so far, $\dot{B}$ and $\dot{B}_s$ are actually negative and we have been plotting them as if they were positive. Thus the correction term in (3) must also be plotted positive. The graphical procedure is then the following:

Table 1. Values of $\dot{\mathscr{I}}B/\mathscr{I}$ at $z_1 = 43$ km as a function of sodium abundance N

N (10^9 atoms/cm^2)	$\dot{\mathscr{I}}B/\mathscr{I}$ (R/km)	N	$\dot{\mathscr{I}}B/\mathscr{I}$
1	0·42	8	16·8
2	1·60	9	19·9
3	3·33	10	23·0
4	5·48	12	29·2
5	8·00	14	35·0
6	10·8	17	43
7	13·7	20	50

plot $\dot{B}$ against $(z_1 + \bar{x})$; plot $\dot{\mathscr{I}}B/\mathscr{I}$ as a zero at $(56 + \bar{x})$ km and the value from Table 1 at $(43 + \bar{x})$ km, drawing a straight line through these points. Add the two curves to give $\dot{B}_s$ which is to be treated as already described in Section 2. The scale of the plot must be known to do this, since the values in Table 1 are given in rayleighs/km. This is easy if an observed B is differentiated numerically to give $\dot{B}$; if the conical-scanning method of Paper VII is used to measure ΔB directly, the scale must be calibrated in some way.

4. DISCUSSION

The methods just described have been used on the observations of nearly a year made with the birefringent photometer (Paper VII) by R. E. Bullock; the results will be reported in another paper. The chord-correction method is easier to apply than the matrix method and also more satisfactory. In a few cases where both methods have been used on the same observations, they have agreed well, though the matrix method gives only a few points instead of a continuous curve. As was mentioned in Section 2, the subtraction technique does not give much improvement because errors of measurement are usually greater than errors of treatment.

It is interesting to notice that a by-product of the chord treatment is a natural definition of a "screening height": the quantity $\bar{x}$ which is between 26 and 27 km. This agrees well with the 25 km found by DUFAY (1947) by taking the point at which $T = 0{\cdot}5$; because $\dot{T}$ is nearly symmetrical this height is approximately equal to $\bar{x}$. Other values of screening height may be found according to the convention used; for example, HUNTEN (1956) obtained –2 km by weighting the transmission function by an exponential to represent the sodium distribution.

This paper has presented a simple method of deriving the vertical distribution of sodium from the derivative of a measured curve of twilight brightness vs. shadow height. This curve is regarded as a smeared version of the distribution and the latter is found by means of the chord construction of BRACEWELL. Tests on computed results were satisfactory and the method has been used on a large number of observations. A modification was also tested; here a computed derivative curve is subtracted before the chord method is applied to the remainder; the result is then added to the assumed initial distribution. This gives a smoother result, but its use on observations we have been able to make does not seem justified. Resonance absorption may have some effect on the shape of the observed derivative; a simple method of correcting this has been developed. Once the abundance is known, a straight line can be drawn on the plot and added to the observations. A series of reliable measurements of vertical distribution may throw some light on the problems of the random and seasonal variations of abundance.

Acknowledgements—Much work in making observations, reducing them, and processing them by the new method has been done by R. E. BULLOCK, B. G. PRATT and J. K. WALKER.

APPENDIX

Normalization of the resonance–absorption correction

We want the value of the term $\dot{\mathscr{S}}B/\mathscr{S}$ in (3) at $z_1 = 43$ km; these are the numbers given in Table 1. The curve representative of the shape of $\dot{\mathscr{S}}$, shown in Fig. 2, is called $s(z_1)$; thus $\dot{\mathscr{S}}(z_1) = \dot{\mathscr{S}}(45)\dot{s}(z_1)$. Similarly, the brightness $B(z_1)$ is written $B_p b(z_1)$ where B_p is the brightness at the plateau. Then

$$\frac{\dot{\mathscr{S}}B}{\mathscr{S}} = \mathscr{S}(45)\,\frac{B_p}{\mathscr{S}}\,.\,\dot{s}b.$$

Now, at the plateau the transmission T is equal to 1; (1) written in full then becomes

$$B_p = k\mathscr{S}\int n\, dz = k\mathscr{S}N$$

where N is the total abundance and k has the value 888 rayleighs/10^9 atoms per cm^2 according to Paper IV (CHAMBERLAIN *et al.*, 1958). Substituting for $B_p/\mathscr{S}$ gives

$$\dot{\mathscr{S}}B/\mathscr{S} = [kN\mathscr{S}(45)]\,.\,(\dot{s}b).$$

From Fig. 2, $\dot{s}b$ is found to be 1·0 at 43 km; it is convenient to tabulate the numbers for this shadow height since they are just equal to $kN\dot{\mathscr{S}}(45)$. $\dot{\mathscr{S}}(45)$ is known as a function of N from the computations of Paper VII.

REFERENCES

BRACEWELL R. N.	1955	*J. Opt. Soc. Amer.* **45,** 873.
BRACEWELL R. N. and ROBERTS J. A.	1954	*Aust. J. Phys.* **7,** 615.
CHAMBERLAIN J. W., HUNTEN D. M. and MACK J. E.	1958	*J. Atmosph. Terr. Phys.* **12,** 153.
DIXON W. R. and AITKEN J. H.	1958	*Canad. J. Phys.* **36,** 1624.
DUFAY J.	1947	*C.R. Acad. Sci., Paris* **225,** 690
HARDY A. C. and YOUNG F. M.	1949	*J. Opt. Soc. Amer.* **39,** 265.
HUNTEN D. M.	1956	*The Airglow and The Aurorae* (Edited by ARMSTRONG E. B. and DALGARNO A.) p. 183. Pergamon Press, London.
RUNDLE H. N., HUNTEN D. M. and CHAMBERLAIN J. W.	1959	*J. Atmosph. Terr. Phys.* **17,** 205.

VOL. 17 59/60

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 302 to 308. Pergamon Press Ltd. Printed in Northern Ireland

A study of the latitude dependence of OH rotational temperatures for Canadian stations*

D. H. McPherson† and A. Vallance Jones
Physics Department, University of Saskatchewan, Saskatoon, Canada

(*Received* 18 *July* 1959)

Abstract—Spectra of the 6,2 and 5,1 OH rotation–vibration bands have been obtained at Saskatoon, Churchill and Resolute Bay over the period of 1957–1958. Rotational temperatures have been derived from the spectra. The mean temperatures obtained at the stations were: Resolute Bay (February–March 1957) 296°K, Saskatoon (February–March, 1957) 227°K, Churchill (March, 1958) 246°K and Saskatoon (April–May, 1958) 227°K. The results were consistent with the latitude effect detected by Chamberlain and Oliver (1953). There appears to be a significant but unexplained difference between the 1957 and 1958 Saskatoon results.

1. Introduction

Since the discovery of the OH rotation–vibrational bands by Meinel (1950a) a number of studies have been made with the object of determining temperatures from the intensity distribution of the rotational lines of the bands. The authors and their results are tabulated in Table 1.

Table 1. Previous determinations of OH temperatures

Author	Latitude	Mean temperature (± range of results)
Meinel (1950b)	42·6°N	260° ± 5°K
Cabannes, Dufay and Dufay (1950)	43·9°N	200° ± 50°K
Dufay and Dufay (1951)	43·9°N	262° ± 7°K
Chamberlain and Oliver (1953)	76·6°N	325° ± 25°K
Gush and Vallance Jones (1955)	52·1°N	200° ± 50°K

The significance of temperatures obtained by this method depends upon whether the excited OH molecules reach rotational equilibrium before radiating. If a value for the average lifetime of 10^{-2} sec is adopted (Heaps and Herzberg, 1952) and the collisional frequency at 90 km is taken to be $1{\cdot}9 \times 10^4$ sec^{-1} (Minzner and Ripley, 1956) then it follows that the number of collisions per lifetime is about 2×10^2 which should be an adequate number for the attainment of thermal equilibrium. It would, therefore, seem fairly safe to conclude that below 90 km the observed band temperature should be equal to the local temperature.

The height of the OH emitting region is not certain. This question has been discussed by Chamberlain and Meinel (1954) who conclude that the emission probably arises in the 70–80 km level. The only rocket data available to date suggest the possibility of a broad emission region extending from 60 to 95 km.

* The Research reported in this paper has been sponsored partly by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19(604)-1831 and partly by the Defence Research Board of Canada under Contract PG69-613013 with the Defence Research Telecommunications Establishment, Ottawa.

† Present address: Department of Applied Mathematics, The Queen's University of Belfast, Belfast, N. Ireland.

VOL. 17 1959/6

KRASSOVSKY (1954, 1957a) argued on theoretical grounds for a height near 100 km. Further rocket results should soon provide a definite answer to this question.

The measured OH rotational temperatures, no doubt, give an average temperature through the emitting region weighted according to the relation between emission intensity and height. Until this intensity–height function is determined, the OH rotational temperatures cannot be compared directly with other temperature measurements.

The purpose of the work described here was primarily to investigate further the latitude effect suggested by the work of CHAMBERLAIN and OLIVER (1953), and of KRASSOVSKY (1957b). As shown in Table 1, CHAMBERLAIN and OLIVER obtained considerably higher rotational temperatures at Thule in Northern Greenland than had been obtained at Yerkes Observatory. A spectrograph suitable for the study of the OH bands had been installed at Resolute Bay by the Defence Research Board and one of us (D. H. McP.) undertook to spend the winter of 1956–1957 using this instrument to obtain OH rotational spectra in the 7000–9000 Å region. During the same winter, plates were obtained at Saskatoon using an optically identical spectrograph. The investigation was extended through to the next winter when plates were exposed at Churchill as well as at Saskatoon. The plates were exposed at Churchill through the kind co-operation of Dr. R. A. Montalbetti of the Canadian Defence Research Board.

The results appear to confirm the existence of a latitude effect and also indicate the occurrence of temperatures as low as 200°K at Saskatoon.

The latitude effect may be due to a reduction of the average height of the emitting layer with increasing latitude if it is assumed that the average height lies in the 60–80 km region. On the other hand, it would not be unreasonable to suppose that there is a relation between the location of the zone of OH production and the temperature profile of the upper atmosphere because both must be related to the basic physical and chemical processes of the system. Variations of OH temperatures with latitude may therefore reflect significant variations in the temperatures of corresponding regions of the atmosphere. Consequently, the OH latitude effect may be an indication of higher mesopheric temperatures in the Arctic as compared with temperate latitudes.

2. EXPERIMENTAL

2.1. *Spectrographs*

The four spectrographs used were of two types. The first type is the 5 in. grating spectrograph described by PETRIE and SMALL (1952). The second type of instrument is larger and employs an 8 in. × 8 in. plane 600 line/mm grating. These spectrographs which are both equipped with $f/0{\cdot}8$ Schmidt cameras, will be referred to as 5 in. and 9 in. spectrographs, respectively. The linear dispersions of these spectrographs were 120 Å/mm and 60 Å/mm, respectively, for the spectra reported here.

2.2. *Observing stations*

The location and equipment of the observing stations are set out in Table 2.

2.3. *Photographic procedure and photometry of spectra*

The spectra were photographed on Kodak I-N plates hypersensitized with an ammonia–ethanol mixture. For each spectral plate the slit was divided into three segments which were exposed as follows: (i) The central segment of the slit was exposed to the night sky at a zenith distance of 70° to 80° in such an azimuth as to minimize contamination of the spectrum by aurora, moonlight or artificial light; exposure varied in length from 3 to 24 hr for the 5 in. spectrograph and up to 60 hr for a 9 in. instrument: (ii) The upper segment of the slit was exposed to an absolutely calibrated low brightness source for 30 min: (iii) The lower segment of the slit was exposed to any bright aurora which occurred.

Table 2. Location and equipment of observing stations

Station	Geographic location		Spectrographs
	Latitude	Longitude	
Saskatoon	52·1°N	106·6°W	5 in. 9 in.
Churchill	58·8°N	94·2°W	9 in.*
Resolute Bay	74·7°N	94·9°W	5 in.*

* The gratings and cameras in these instruments are of exactly the same type as those of the Saskatoon spectrographs.

VOL. 17 1959/6

A calibration spectrum was obtained by exposing again for 30 min to the low brightness source with a rotating step sector placed at the slit. The exposure was made on the same plate as for the nightglow spectrum when the 9 in. spectrographs were used, but had to be made on a separate plate with the 5 in. instruments. Both 5 in. spectrograph plates were always developed together.

After development the regions of the OH bands on the night sky plates were microphotometered and the results reduced to give the density across the spectrum relative to the clear plate density. From the step sector calibration, characteristic curves were prepared for all wavelength points at which it was necessary to measure the relative brightness of the spectrum. Finally, it was necessary to subtract the contributions of any background continuum which is always present even on the cleanest plates, and seems to arise mostly from two sources; (i) scattered moonlight and (ii) scattered tungsten light. The relative contributions of the two are indicated by the occurrence and central "depth" of the Fraunhöfer *A* band. Spectra of scattered sunlight were measured with the spectrographs using rotating step sectors at the slits and the same spectrograph settings as for the night sky exposures. The scattered tungsten light was assumed to follow the emission of a 2900°K black body. When the contribution of the two types of continuum had been estimated from an inspection of the uncorrected spectrum, this continuum correction was subtracted to yield corrected values for the peak brightness of the rotational lines of the *P*-branch.

3. Results

3.1. *Derivation of rotational temperatures*

The number of plates suitable for measurement at each station during the period of the investigation is indicated in Table 3. The relative peak intensities of the $P_1(3/2)$, $P_1(5/2)$, $P_1(7/2)$, $P_1(9/2)$ and $P_1(11/2)$ lines were obtained from these plates for the 6,2 OH bands and in several cases for the 5,1 bands as well.

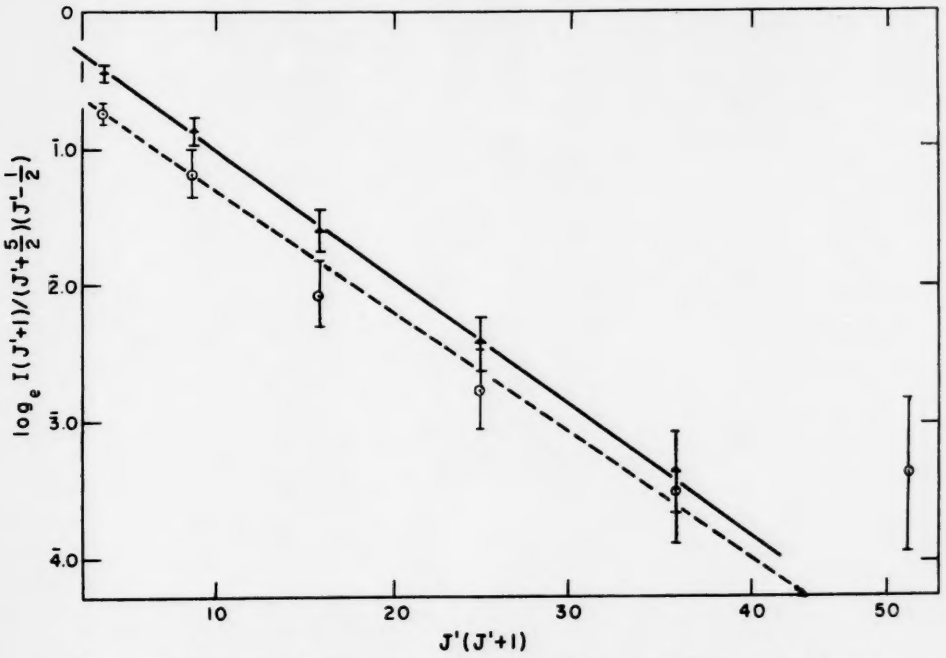


Fig. 1. Plots of $\ln I\{P(J'')\}\ (J' + 1)/(J' + 5/2)(J' - 1/2)$ against $J'(J' + 1)$ for best and worst cases at Saskatoon. ▲ (solid curve), best case; ⊙ (dotted curve), worst case.

Rotational temperatures were determined from these data with the help of the theoretical line intensity relation discussed by Meinel (1950b). As pointed out by Meinel, the assumption of Hund's case (a) coupling should be valid for the J-values considered. The value of B_v was calculated taking $B_e = 18{\cdot}86$, $\alpha = 0{\cdot}69$ and $\Omega = 3/2$ for the P_1-branch. Consequently, one may obtain the temperature from the slope of a plot of

$$\log\left[I\,\frac{\{(P(J'')\}\,.\,(J' + 1)}{(J' + 5/2)(J' - 1/2)}\right]$$

against $J'(J' + 1)$. Two examples of the plots obtained are shown in Fig. 1; these examples correspond to the best and worst plots obtained from Saskatoon.

The collected data for relative P-branch line intensities and derived rotational temperatures are set out in Table 3.

Table 3. A summarized table of observations and results

Observing station	Year	Exposure period	Spectrograph	Bands measured	$I\{P_1(J'')\}/I\{P_1(3/2)\}$ J'' 3/2	5/2	7/2	9/2	11/2	13/2	Temperature °K	Remarks
Saskatoon	1956	Nov. 22–25	9 in.	(6,2)	1·00	1·28	1·10	0·68	0·40	—	286 ± 20	
	1957	Feb. 24–25	5 in.	(6,2)	1·00	1·19	0·98	0·66	0·33	—	273 ± 10	
		Mar. 14–17	5 in.	(6,2)	1·00	1·21	0·97	0·75	0·34	0·22	286 ± 10	
		Mar. 17–23	5 in.	(6,2)	1·00	1·26	0·84	0·48	0·20	0·18	232 ± 20	
	1958	Jan. 29–Feb. 6	9 in.	(6,2)	1·00	1·27	0·76	0·82	0·25	—	269 ± 80	Under exposed.
		Feb. 7–23	9 in.	(6,2)	1·00	1·09	0·80	0·56	″—	—	244 ± 50	
		Apr. 1	5 in.	(6,2)	1·00	1·36	0·83	0·38	0·28	—	224 ± 20	
		Apr. 11–13	5 in.	(6,2)	1·00	1·17	0·68	0·45	0·21	0·17	221 ± 15	
		Apr. 19–22	5 in.	(6,2)	1·00	1·34	0·89	0·44	0·17	0·094	211 ± 15	Two spectra on one plate. Both (6,2) and (5,1) bands measurable
				(5,1)	1·00	1·10	0·85	0·37	0·19	—	253 ± 25	
		Apr. 23–24	5 in.	(6,2)	1·00	1·17	0·86	0·37	0·11	0·062	189 ± 10	
			5 in.	(5,1)	1·00	1·22	0·90	0·37	0·15	—	258 ± 25	
		May 17	5 in.	(6,2)	1·00	1·19	0·70	0·39	0·28	—	198 ± 20	Two spectra on one plate. Both (6,2) and (5,1) measurable.
			5 in.	(5,1)	1·00	1·13	0·89	0·54	0·38	0·11	210 ± 30	
		May 18–19	5 in.	(6,2)	1·00	1·14	0·66	0·42	0·24	0·32	230 ± 20	
			5 in.	(5,1)	1·00	1·11	0·88	0·54	0·38	0·12	215 ± 10	
Churchill	1958	Mar. 4–8	9 in.	(6,2)	1·00	1·11	0·83	0·62	0·37	0·23	247 ± 20	
		Mar. 25–27	9 in.	(6,2)	1·00	1·19	0·89	0·57	—	—	245 ± 20	
Resolute Bay	1957	Feb. 27–Mar. 1	5 in.	(6,2)	1·00	1·35	1·04	0·74	0·55	—	263 ± 10	
		Mar. 4–7	5 in.	(6,2)	1·00	1·34	0·94	0·74	0·38	—	271 ± 20	
		Mar. 17–19	5 in.	(6,2)	1·00	1·32	1·10	0·81	—	—	315 ± 20	Measured twice.
			5 in.	(6,2)	1·00	1·23	1·07	0·75	0·67	—	297 ± 20	
		Mar. 20–23	5 in.	(6,2)	1·00	1·27	1·11	0·97	—	—	346 ± 15	Measured twice.
			5 in.	(6,2)	1·00	1·43	1·29	1·09	0·46	—	345 ± 15	

VOL. 17 1959/6

4. Discussion of Results

The average temperatures calculated from the results given in Table 3 are collected in Table 4. The results have been grouped together in various ways in order to examine the existence of any latitude of time variations.

The results of Table 4 can be interpreted in two ways: first, if one compares the Saskatoon average temperature with the Resolute Bay average, one finds that the difference between the two averages is highly significant (on the basis of Student's *t*-test for small samples). This would lead to the conclusion that the OH rotational temperatures exhibit a latitude effect. The intermediate value of the

Table 4. Average OH temperatures

Station	Year	No. of observations	Mean temperature (°K)	Standard deviation
Saskatoon	1958	12	227	±25
Saskatoon	1957	3	264	±23
Saskatoon	1957–1958	15	234	±29
Churchill	1958	2	246	
Resolute Bay	1957	4	296	±38

OL.
17
9/60

average temperature obtained at Churchill would be consistent with this interpretation, even though it is not itself significant. Secondly, however, if one subdivides the Saskatoon results into two groups for 1957 and 1958 one finds a significant difference between the averages for these groups; at the same time the possibility that the difference between the Saskatoon and Resolute Bay averages for 1957 is due to chance, rises to 25 per cent. Our results for 1957, therefore, do not definitely establish the existence of the latitude effect but they do provide strong support for its existence. The significance of the difference between the 1957 and 1958 Saskatoon temperatures is not clear. It could arise from a variation of OH rotational temperature with solar activity or from a seasonal variation, since Table 3 shows that three of the four 1957 spectra were obtained in February and March while nine out of eleven 1958 spectra were obtained in April and May. Further work will be required to distinguish between these possibilities.

An important result of this study is to show that OH rotational temperatures fall anywhere between 200°K and 300°K at Saskatoon, and indeed were frequently observed in the 200–220°K range in April and May 1958. The temperature of 200°K deduced by Gush and Vallance Jones (1953), for the rotational temperature of the 4,2 OH band thus appears reasonable, especially since this spectrum was obtained in June.

5. Conclusions

The comparison between the OH rotational temperatures observed at Resolute Bay, Churchill and Saskatoon is consistent with the existence of a latitude effect as suggested by the work of Chamberlain and Oliver (1953). The observed difference in average temperatures for 1957 amounted to 32°K. On the other hand, a comparison of the results obtained at Saskatoon in February–March 1957 and

April–May 1958 shows the existence of a 37°K difference in mean temperatures between the two sets of observations. This difference is statistically significant but it is not clear whether the variation is a seasonal one or some other type. Further measurements of OH rotational temperatures using a more rapid and accurate method than the photographic one would be very desirable.

Acknowledgements—We should like to acknowledge the help of Mr. H. J. Koenig for operation of the Saskatoon spectrograph and of Dr. R. A. Montalbetti, Officer in charge of the Defence Research Northern Laboratory of the Defence Research Board of Canada for the co-operation of himself and staff in obtaining the spectra at Churchill.

References

Cabannes J., Dufay J. and Dufay M.	1950	*C.R. Acad. Sci., Paris* **230,** 1233.
Chamberlain J. W. and Meinel A. B.	1954	*The Earth as a Planet* (Edited by Kuiper, G. P.) p. 514. University of Chicago Press, Chicago.
Chamberlain J. W. and Oliver N. J.	1953	*Phys. Rev.* **90,** 1118.
Dufay J. and Dufay M.	1951	*C.R. Acad. Sci., Paris* **232,** 426.
Gush H. P. and Vallance Jones A.	1953	*J. Atmosph. Terr. Phys.* **7,** 285.
Heaps H. S. and Herzberg G.	1952	*Z. Phys.* **133,** 48.
Krassovsky V. I.	1954	*Usp. Fiz. Nauk.* **54,** 469.
Krassovsky V. I.	1957a	*J. Atmosph. Terr. Phys.* **10,** 49.
Krassovsky V. I.	1957b	*Priroda* **12,** 87.
Meinel A. B.	1950a	*Astrophys. J.* **111,** 555.
Meinel A. B.	1950b	*Astrophys. J.* **112,** 120.
Minzner R. A. and Ripley W. S.	1956	*Air Force Surveys in Geophysics,* TN-56-204, Dec. No. 86.
Petrie W. and Small R.	1952	*Astrophys. J.* **116,** 433.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 309 to 314. Pergamon Press Ltd. Printed in Northern Ireland

Group refractive index of the ionosphere at low frequency

Y. S. N. Murty and S. R. Khastgir
Physics Department, University College of Science, Calcutta 9

(*Received* 27 *July* 1959)

Abstract—A general expression is obtained for the group refractive index for low-frequency radio wave propagation in the ionosphere. It is shown that, in the limiting case when electron collisional frequency is small, this expression for the group refractive index is identical with the approximate expression obtained by Gibbons and Rao.

Introduction

Taking the electron-density profile in the ionosphere along the Z-direction, the expression for the group height is given by:

$$\int \mu' \, dz = \int \left(\mu + f \frac{\partial \mu}{\partial f} \right) dz \tag{1}$$

where μ' is the group refractive index, μ the phase refractive index and f the frequency of the exploring wave. For obtaining the group height for a given electron-density profile, it is necessary to know the value of the group refractive index for different values of electron-density and of electron-collisional frequency. There are two methods of determining the group refractive index. In one method, the quantity $\partial\mu/\partial f$ is evaluated by computing μ at two closely spaced frequencies on either side of the operating freqeuncy using the Appleton–Hartree magneto-ionic formulae (Appleton, 1927; Hartree, 1929). In the other method, the Appleton–Hartree equation for the complex refractive index is differentiated to obtain an analytical expression for $\partial\mu/\partial f$. In either method the group refractive index μ' is then obtained from (1).

Gibbons and Rao (1957) obtained an analytical expression for μ' by making an approximation in the Appleton–Hartree formulae which is applicable to the Pennsylvania State University area, for which the critical collisional frequency ν_e is small. In the present communication, an analytical expression for μ' is obtained with no such approximation in the Appleton–Hartree formulae. In the limiting case when the electron–collisional frequency ν_e is small, this expression is shown to be the same as that given by Gibbons and Rao (1957).

Deduction of the Expression for the Group Refractive Index

From the Appleton–Hartree formulae we can write,

$$M^2 - 1 = \frac{1}{\alpha + i(\beta + R\gamma_L)} \tag{2}$$

and

$$R = \frac{1}{i\gamma_L}\left[- \frac{{\gamma_T}^2}{2(1 + \alpha + i\beta)} \pm \sqrt{\left(\frac{{\gamma_T}^4}{4(1 + \alpha + i\beta)^2} + {\gamma_L}^2 \right)} \right] \tag{3}$$

where M = complex refractive index;
R = complex polarization;

$$\alpha = -\frac{p^2}{{p_0}^2}, \qquad \beta = \frac{p\nu}{{p_0}^2}, \qquad p_H = \frac{eH}{mc},$$

$$\gamma_L = \frac{pp_H \cos\theta'}{{p_0}^2}, \qquad \gamma_T = \frac{pp_H \sin\theta'}{{p_0}^2}, \qquad {p_0}^2 = \frac{4\pi Ne^2}{m}$$

and θ' = angle between the direction of propagation and the earth's magnetic field.

(The other notations have their usual significance).

From equation (1) we have,

$$\mu' - \mu = f\frac{\partial\mu}{\partial f} = R'\left(f\frac{\partial M}{\partial f}\right) \tag{4}$$

where R' denotes "real part of".

Differentiating equation (2) with respect to f, we get

$$2fM\frac{\partial M}{\partial f} = -(M^2-1)^2 f\frac{\partial}{\partial f}(\alpha + i\beta + iR\gamma_L)$$
$$= 1 - M^2 - (M^2-1)^2\left(\alpha + i\gamma_L f\frac{\partial R}{\partial f}\right). \tag{5}$$

Putting $M = (\mu - i\chi)$, where $\chi = ck/p$, k = absorption index, $p/2\pi$ = frequency of the exploring wave and c = velocity of light, and multiplying both sides of equation (5) with the complex conjugate of M, we have

$$2f\frac{\partial M}{\partial f}(\mu^2+\chi^2) = (\mu + i\chi)\Big[1 - \mu^2 + \chi^2 + 2i\mu\chi$$
$$- \left\{\alpha + i\gamma_L f\frac{\partial R}{\partial f}\right\}\left\{(\mu^2 - \chi^2 - 1)^2 - 4\mu^2\chi^2 - 4i\mu\chi(\mu^2 - \chi^2 - 1)\right\}\Big].$$

Writing $\tau = 1 + \chi^2 - \mu^2$

and $R = \rho e^{i\phi} = \rho\cos\phi + i\rho\sin\phi = r + is$

where ρ is the ratio of the normal to the abnormal component of the magnetic vector of the wave, ϕ the phase-difference between them and $r = \rho\cos\phi, s = \rho\sin\phi$, we obtain,

$$2f\frac{\partial M}{\partial f}(\mu^2+\chi^2) = (\mu + i\chi)\Big[\tau + 2i\mu\chi$$
$$- \left\{\tau^2 - 4\mu^2\chi^2 + 4i\mu\chi\tau\right\}\left\{\alpha + i\gamma_L f\left(\frac{\partial r}{\partial f} + i\frac{\partial s}{\partial f}\right)\right\}\Big]. \tag{6}$$

Taking the real parts on both sides of equation (6), we get

$$2(\mu' - \mu)(\mu^2+\chi^2) = \mu(1 - \mu^2 - \chi^2) - \mu\tau^2\left(\alpha - \gamma_L f\frac{\partial s}{\partial f}\right)$$
$$+ 4\mu\chi^2(1+\chi^2)\left(\alpha - \gamma_L f\frac{\partial s}{\partial f}\right) + \chi\gamma_L f\frac{\partial r}{\partial f}\left\{\tau^2 + 4\mu^2(1-\mu^2)\right\}. \tag{7}$$

VOL. 17 1959/6

The quantity $f\dfrac{\partial s}{\partial f}$ in equation (7) is obtained in the following way.

Following the Appleton–Hartree formulae it was shown elsewhere* (MURTY, 1959) that the ratio of the normal to the abnormal component of the magnetic vector for the ordinary wave† is given by

$$\rho = \frac{a'}{\nu_c}\left[\frac{\nu}{\cos\phi} - \frac{p'}{\sin\phi}\right] \tag{8}$$

and the phase-difference ϕ is given by

$$\cos 2\phi = a' - \sqrt{\{(1+a')^2 - 4a'b'\}} \tag{9}$$

where $a' = \dfrac{\nu_c^2}{\nu^2 + p'^2}$, $\quad b' = \dfrac{\nu^2}{\nu^2 + p'^2}$ and $\quad p' = p\left(1 - \dfrac{p_0^2}{p^2}\right)$.

Since $s = \rho \sin\phi = \dfrac{a'}{\nu_c}(\nu \tan\phi - p')$ (10)

and

$$a' = \frac{\nu_c^2}{\nu^2 + p'^2}. \tag{11}$$

it can be very easily shown that

$$\frac{\partial s}{\partial a'} = \frac{s}{a'} - \frac{a'\nu}{\nu_c}\cdot\frac{\sec^2\phi}{2\sin 2\phi}\left[1 - \frac{(1+a') - 4\dfrac{\nu^2}{\nu_c^2}a'}{\sqrt{\left\{(1+a')^2 - 4\dfrac{\nu^2}{\nu_c^2}\cdot a'^2\right\}}}\right] + \frac{\nu_c}{2a'p'} \tag{12}$$

and

$$f\frac{\partial a'}{\partial f} = -\frac{2p'\nu_c^2}{(\nu^2 + p'^2)^2}\left[p + \frac{p_0^2}{p}\right] = -\frac{2a'^2}{\nu_c^2}pp'\left(1 + \frac{p_0^2}{p^2}\right). \tag{13}$$

Therefore

$$f\frac{\partial s}{\partial f} = \left(\frac{\partial s}{\partial a'}\right)\left(f\frac{\partial a'}{\partial f}\right) = -\frac{2a'^2}{\nu_c^2}pp'\left(1 + \frac{p_0^2}{p^2}\right)\left[\frac{s}{a'} + \frac{\nu_c}{2a'p'} - \frac{\dfrac{a'\nu}{\nu_c}\left\{1 + \dfrac{(1+a') - 4b'}{\sqrt{\{(1+a')^2 - 4a'b'\}}}\right\}}{[(1+a') - \sqrt{\{(1+a')^2 - 4a'b'\}}][2a'(2b' - 1 - a' + \sqrt{\{(1+a')^2 - 4a'b'\}})]}\right] \tag{14}$$

Thus for places of low critical collisional frequency and for low frequencies of the exploring wave, $f \simeq 150$ ks/s, $\dfrac{p_H}{p} \simeq 10$, $\dfrac{\nu_c}{p} \simeq 0{\cdot}5$, and $\nu \simeq 10^6$, we can write‡

$$f\frac{\partial s}{\partial f} \simeq -\frac{2pp'a'^2}{\nu_c^2}\left(1 + \frac{p_0^2}{p^2}\right)\left(\frac{\nu^2}{2p'\nu_c}\right) \simeq -\frac{a'}{\nu_c}p\left(1 - \frac{1}{\alpha}\right).$$

---

* *Vide* Appendix I.

† It is to be noted that we are concerned with only the ordinary wave, as the extraordinary wave is completely absorbed in the ionosphere at low frequencies.

‡ *Vide* Appendix II.

Since $r = \rho \cos \phi = \frac{a'}{v_c}(v - p' \cot \phi) \simeq \frac{a'v}{v_c}$, we have

$$f \frac{\partial s}{\partial f} \simeq -\frac{r}{v} p\left(1 - \frac{1}{\alpha}\right) \tag{15}$$

Substituting this value of $f \frac{\partial s}{\partial f}$ in the term $\mu\tau^2\gamma_L f \frac{\partial s}{\partial f}$ in (7), we have

$$2(\mu' - \mu)(\mu^2 + \chi^2) = \mu(1 - \mu^2 - \chi^2) - \mu\tau^2\left\{\alpha + \gamma_L \frac{r}{v} p\left(1 - \frac{1}{\alpha}\right)\right\} + \\ + 4\mu\chi^2(1 + \chi^2)\left(\alpha - \gamma_L f \frac{\partial s}{\partial f}\right) + \chi\gamma_L f \frac{\partial r}{\partial f}(\tau^2 + 4\mu^2\sqrt{1 - \mu^2}). \tag{16}$$

Dividing throughout by $2\mu(\mu^2 + \chi^2)$ we have,

$$\frac{\mu'}{\mu} = 0{\cdot}5 + \frac{1}{\mu^2 + \chi^2}\left[0{\cdot}5 - \frac{\tau^2}{2}\left\{\alpha + \gamma_L \frac{\tau}{v} p\left(1 - \frac{1}{\alpha}\right)\right\} + \right. \\ \left. + 2\chi^2(1 + \chi^2)\left(\alpha - \gamma_L f \frac{\partial s}{\partial f}\right) + \frac{\tau^2}{2}\chi\gamma_L f \frac{\partial r}{\partial f}\left\{\frac{1}{\mu} + \frac{4\mu}{\tau^2}(1 - \mu^2)\right\}\right]. \tag{17}$$

It can be easily shown by numerical computation that,

$$\frac{1}{\mu^2 + \chi^2}\left[2\chi^2(1 + \chi^2)\left(\alpha - \gamma_L f \frac{\partial s}{\partial f}\right) + \frac{\tau^2}{2}\chi\gamma_L f \frac{\partial r}{\partial f}\left\{\frac{1}{\mu} + \frac{4\mu}{\tau^2}(1 - \mu^2)\right\}\right] \simeq 0$$

so that

$$\frac{\mu'}{\mu} = 0{\cdot}5 + \frac{1}{\mu^2 + \chi^2}\left[0{\cdot}5 - \frac{\tau^2}{2}\left\{\alpha + \gamma_L \frac{r}{v} p\left(1 - \frac{1}{\alpha}\right)\right\}\right].$$

Writing $\frac{pp_H}{{p_0}^2} \cos \phi'$ for γ_L, we have

$$\frac{\mu'}{\mu} = 0{\cdot}5 + \frac{1}{\mu^2 + \chi^2}\left[0{\cdot}5 - \frac{\tau^2}{2}\left\{\alpha + (1 - \alpha)\, p_H \cos \phi' \frac{r}{v}\right\}\right]. \tag{18}$$

Discussion of the Results of Gibbons and Rao

Gibbons and Rao (1957) obtained the approximate expression for $\frac{\mu'}{\mu}$ as:

$$\frac{\mu'}{\mu} = 0{\cdot}5 + \frac{1}{\mu^2 + \chi^2}\left[0{\cdot}5 + \frac{\tau^2}{2}\left\{-\alpha + (1 - \alpha)p_H \cos \phi' . \frac{r}{v}\right\}\right]. \tag{19}$$

On comparing (18) and (19) it is evident that the term

$$\frac{1}{\mu^2 + \chi^2} \cdot \frac{\tau^2}{2}(1 - \alpha)p_H \cos \phi' \, . \frac{r}{v}$$

in the expression given by Gibbons and Rao has a positive sign while in the expression as given in (18) this term has a negative sign.

VOL. 17 1959/6

On scrutinizing the expression of GIBBONS and RAO, viz.

$$f\frac{\partial s}{\partial f} = (1+\alpha')\frac{r}{\beta} - \frac{2\beta(1+\alpha')(2r-\beta\epsilon)}{(1-\alpha')^2+\beta^2} \tag{20}$$

where $\alpha' = -\alpha = \dfrac{p^2}{{p_0}^2}$ and $\epsilon = \dfrac{p_H}{p\sigma}\cdot\dfrac{\alpha'}{(1-\alpha')^2+\beta^2}$

it can be easily shown that

$$f\frac{\partial s}{\partial f} \simeq -\frac{2pp'a'^2}{{\nu_c}^2}\left(1-\frac{1}{\alpha}\right)\left(\frac{\nu^2}{2p'\nu_c}\right) \simeq -\left(1-\frac{1}{\alpha}\right)\frac{pa'}{\nu_c} \simeq -\frac{r}{\nu}p\left(1-\frac{1}{\alpha}\right).$$

This is the same as given in (15).

It is to be noted that, due to neglecting the subtracting term in (20), GIBBONS and RAO obtained,

$$f\frac{\partial s}{\partial f} = (1+\alpha')\frac{r}{\beta} = +\frac{r}{\nu}p\left(1-\frac{1}{\alpha}\right).$$

Accepting the expression for $f\,(\partial s/\partial f)$ as given in (15) it is evident that the value of μ'/μ as obtained in (18) should be less than the value obtained from the corresponding formula given by GIBBONS and RAO. This explains why GIBBONS and RAO obtained from their formula values of μ' higher than those obtained by the laborious but accurate method of computing μ' from values of μ at two closely spaced frequencies on either side of the operating frequency.

Acknowledgement—We are thankful to Mr. S. DATTA MAJUMDAR for valuable suggestions.

REFERENCES

APPLETON E. V.	1927	*URSI Papers*, Washington.
GIBBONS J. J. and RAO B. R.	1957	*J. Atmosph. Terr. Phys.* **11,** 151.
HARTREE D. R.	1929	*Proc. Camb. Phil. Soc.* **25,** 47.
MURTY Y. S. N.	1959	*Sci. & Cult.* **25,** 161.

APPENDIX I

Denoting the value of R with negative and positive signs before the radical in (3) as R_0 and R_x corresponding to the ordinary and the extraordinary components we have

$$R_0 + R_x = \frac{i{\gamma_T}^2/\gamma_L}{(1+\alpha+i\beta)} = \frac{2}{x+iy} \tag{21}$$

and $R_0 \,.\, R_x = 1$ (22)

where $x = \dfrac{\sigma\nu}{p_H}$, $\quad y = \dfrac{\sigma p'}{p_H}$, $\quad \sigma = \dfrac{2\cos\phi'}{\sin^2\phi'}$, $\quad p' = p\left(1-\dfrac{{p_0}^2}{p^2}\right)$.

Writing $R = \rho e^{i\phi}$ we get from (21) and (22)

$$\rho e^{i\phi} + \frac{1}{\rho}e^{-i\phi} = \frac{2}{x+iy} = 2a'(x-iy) \tag{23}$$

where $a' = \nu_c^2/(\nu^2+p'^2)$ = critical collisional frequency.

Equating real and imaginary parts in (23) we obtain

$$\left(\rho+\frac{1}{\rho}\right)\cos\phi = 2a'x = \frac{2a'\nu}{\nu_c} \tag{24}$$

and
$$\left(\rho - \frac{1}{\rho}\right) \sin\phi = -2a'y = -\frac{2a'p'}{v_c}. \tag{25}$$

From (24) and (25) we have

$$\rho = \frac{a'}{v_c}\left[\frac{v}{\cos\phi} - \frac{p'}{\sin\phi}\right] \tag{26}$$

and
$$\rho^{-1} = \frac{a'}{v_c}\left[\frac{v}{\cos\phi} + \frac{p'}{\sin\phi}\right]. \tag{27}$$

On eliminating ρ from (26) and (27) we get

$$\cos 2\phi = a' - \sqrt{\{(1+a')^2 - 4a'b'\}} \tag{28}$$

where
$$b' = \frac{v^2}{v^2 + p'^2}.$$

Appendix II

Equation (14) is as follows:

$$f\frac{\partial s}{\partial f} = -\frac{2a'^2}{v_c^2} pp' \left(1 + \frac{p_0^2}{p^2}\right)\left[\frac{s}{a'} + \frac{v_c}{2a'p'} - \right.$$

$$\left. - \frac{\dfrac{a'v}{v_c}\left\{1 - \dfrac{(1+a') - 4b'}{\sqrt{\{(1+a')^2 - 4a'b'\}}}\right\}}{[(1+a') - \sqrt{\{(1+a')^2 - 4a'b'\}}]\,[2a'(2b' - 1 - a' + \sqrt{\{(1+a')^2 - 4a'b'\}})]^{\frac{1}{2}}}\right]$$

For $p'^2 \ll v^2$, $b' \simeq 1$ and the last term (within the square bracket) of the above equation can be written as

$$\frac{1}{(1+a') - (1-a')} \times \frac{1}{[2a'(1-a'+1-a')]^{\frac{1}{2}}} \times \frac{a'v}{v_c}\left(1 - \frac{a'-3}{1-a'}\right).$$

This may be further simplified as

$$\frac{1}{2a'^{\frac{1}{2}}(1-a')^{\frac{1}{2}}} \times \frac{v}{v_c} \times \left(\frac{2-a'}{1-a'}\right).$$

This ultimately is reduced to

$$\frac{1}{2a'}\left[2 + \frac{a'}{1-a'}\right] \simeq \frac{1}{a'} + \frac{1}{2}\left(\text{since } a' \simeq \frac{v_c^2}{v^2}\right).$$

Since
$$s = \rho \sin\phi = \frac{a'}{v_c}(v\tan\phi - p') \simeq \frac{a'}{v_c} v\tan\phi$$

it can be shown that

$$\frac{s}{a'} = \frac{v}{v_c}\tan\phi = \frac{1}{a'} - \frac{1}{2}.$$

Hence

$$f\frac{\partial s}{\partial f} \simeq -\frac{2pp'a'^2}{v_c^2}\left(1 + \frac{p_0^2}{p^2}\right)\left(\frac{v^2}{2p'v_c} - 1\right) \simeq -\frac{a'}{v_c} p\left(1 - \frac{1}{\alpha}\right) \simeq -\frac{r}{v} p\left(1 - \frac{1}{\alpha}\right)$$

$$\left(\text{since } \frac{v^2}{2p'v_c} \gg 1\right).$$

VOL. 17 1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 315 to 319. Pergamon Press Ltd. Printed in Northern Ireland

A relationship between the life of *M*-regions and the rate of change of solar activity

D. W. G. CHAPPELL
Dominion Physical Laboratory, Lower Hutt, New Zealand

(*Received* 28 *July* 1959)

Abstract—To test the hypothesis that the observed lack of success in forecasting disturbances to radio propagation by using the recurrence pattern during the 1953–1954 sunspot minimum period is due to the high rate of decay of the 1947–1954 cycle of solar activity, magnetic character figures for the period 1890–1950 were examined.

A numerical definition of "persistence of recurrence" s given and a linear relationship is obtained between the maximum value of this quantity at sunspot minimum and the rate of decay of solar activity to that minimum.

The extremely low value of this quantity predicted by this relationship for the 1954 sunspot minimum period is considered to provide confirmatory evidence for the initial hypothesis.

INTRODUCTION

THE CHAPMAN–FERRARO (1931–1933) theory of ionospheric and magnetic storms postulates that they result from the interaction between the ionosphere and a neutral stream of charged particles emanating from the sun. Two well-defined types of emission have been recognized, an impulsive type which results from the solar eruptions associated with large, active sunspots and causes large, non-recurrent storms with sudden commencements; and a steady type, whose source is not clearly understood, which causes small, recurrent and slow-starting storms. As would be expected the former type is prevalent in sunspot maximum years and the latter in sunspot minimum years.

To explain this second type of emission, BARTELS (1932) has postulated the existence of especially active areas on the sun (called *M*-regions) which emit particles continuously throughout their lifetime. As the sun rotates with a mean synodic period of $27\frac{1}{4}$ days, these particles form a rotating spiral in space which will overtake the earth once in each solar rotation resulting in the observed recurrence of the associated storms. A number of tentative hypotheses have been proposed to explain the physical nature of these *M*-regions (*inter alia* WALDMEIER, 1946; ALLEN, 1944; and BABCOCK and BABCOCK, 1955), but none has so far achieved general acceptance.

As long distance radio propagation depends for its success on the ionosphere, an ionospheric storm will result in a radio propagational disturbance. Such disturbances were forecast with considerable success during the sunspot minimum period of 1943–1945, but this success was not repeated during the 1954 sunspot minimum period.

This observation could well be the result of the high rate of decay of the 1947–1954 cycle of solar activity, for it could be expected that the higher the rate of change of solar activity the shorter will be the life of any solar phenomenon, in particular *M*-regions. This investigation was undertaken to discover whether there is any evidence to support this hypothesis.

DATA

As M-regions are not directly observable, a study of their life must devolve into a study of the tendency to recur of one of their terrestrial manifestations (*inter alia* ionospheric storms, magnetic storms, radio propagational disturbances). Of these, magnetic storms provide the most promising avenue of enquiry for they alone have been observed for more than about 15 years and furthermore only these data are available in a form directly suitable for analysis.

Each of the world-wide chain of magnetic observatories (some 50 in 1890 and some 100 in 1940, CHAPMAN and BARTELS, 1940), inspects its records of the diurnal variation of the magnetic elements and places the day into one of three categories; 0 if the records represent a quiet day, 1 if a moderately disturbed day, or 2 if a considerably disturbed day. Zurich Observatory determines the mean value of these figures to one place of decimals and publishes it as the International Magnetic Character Figure for the day (C-figures). These C-figures are undoubtedly subjective but, being the mean value of over fifty independent judgements, could be expected to exhibit most of the properties of objective data: however, there will be systematic differences between C-figures recorded at times separated by many years, but reasonable homogeneity should be obtained over a period of 3 years which is all that is required by the analysis.

By the choice of a suitable threshold value, these figures can be used to classify days into two groups—quiet and storm days. The value chosen for this purpose was $C = 0{\cdot}7$ for two reasons; firstly, it is the intermediate value chosen by BARTELS (1932) to separate his red and black symbols, and secondly, this value results in a relatively large number of storm days enabling the analysis to be carried out using large numbers, thereby reducing the influence of random variations. Thus a day will be considered a storm day if its C-figure exceeds 0·7.

VOL. 17 1959/6

DEFINITION OF PERSISTENCE

If in any period of N days we have n_1 storm days and of these n_2 have storm days 27 days later, we wish to combine these three numbers in such a way that the resultant figure will represent the tendency of storms to recur during the period.

More generally, let there be a non-reciprocal law associating every day with some other day and in a period of N days there are n_1 days exhibiting a given character and n_2 pairs of associated days, both of which exhibit this character. If there is no relationship between the chosen law of association and the given character, then it can be shown by an argument precisely similar to that used by MOOD (1940) in his theory of random runs that, neglecting end effects, the expected value of the ratio:

$$\frac{n_2}{n_1} \quad \text{is} \quad \frac{n_1}{N}.$$

If, on the other hand, there is a tendency for days exhibiting the given character to be associated by the chosen law, then the ratio n_2/n_1 will exceed n_1/N by a greater or less extent as the tendency to be associated is greater or less.

In this case the character is "being a storm day" and the law is "that the day

associated with day D is day $D + 27$", so we can define the recurrence tendency for the period or persistance (P), as it will be called, by

$$P = \frac{n_2}{n_1} - \frac{n_1}{N}$$

where N is the number of days if the period, n_1 is the number of storm days, and n_2 the number of pairs of storm days separated by 27 days contained therein.

Method of Analysis

(a) *Determination of the persistence of recurrence*

The period 1890–1957 was divided into running 3 year periods and, for each, the values of N, n_1 and n_2 were counted and P calculated. The diagram given by Bartels (1932) was used for the period 1907–1929, similar diagrams were constructed for the rest of the period.

(b) *Determination of the rate of decay of solar activity*

It has been established for many years that the time variation of the running mean of the sunspot number is closely related to that of solar activity. As in the

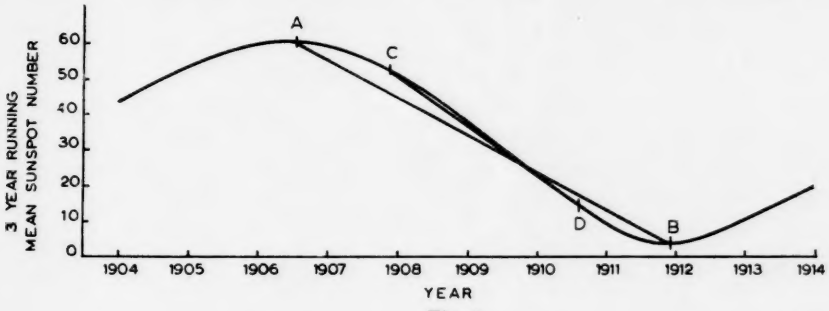


Fig. 1.

determination of the values for persistence of recurrence a 3 year period was used a similar period was taken for the calculation of these running mean values.

At first glance it would seem that an acceptable value for the rate of decay of solar activity would be the slope of the line joining points representing the maximum and minimum values (Fig. 1, points A and B). However, this value is seriously affected by curvature near the critical points of the curve; the points C and D would appear to give a more acceptable value for the rate of decay of solar activity. In this work the time intervals AC, BD are one-quarter that of AB.

Results

The time variation of P and sunspot number are shown in Fig. 2 and, as expected, they are approximately in antiphase. In Fig. 3 are related the maximum values of the P curve and the rate of decay of solar activity towards the appropriate minimum. It will be seen that, in spite of the arbitrary nature of the definitions of these quantities, the points lie remarkably close to a straight line. This being so

it can be inferred that there is a tendency for the maximum value of the persistence of recurrence of magnetic storms over a sunspot minimum period to vary in an inverse manner with the rate of decay of solar activity to that minimum and further that the length of life of M-regions will likewise vary inversely with that rate of decay.

This result provides confirmatory evidence of the initial hypothesis that the observed lack of success in forecasting radio propagational disturbances using the

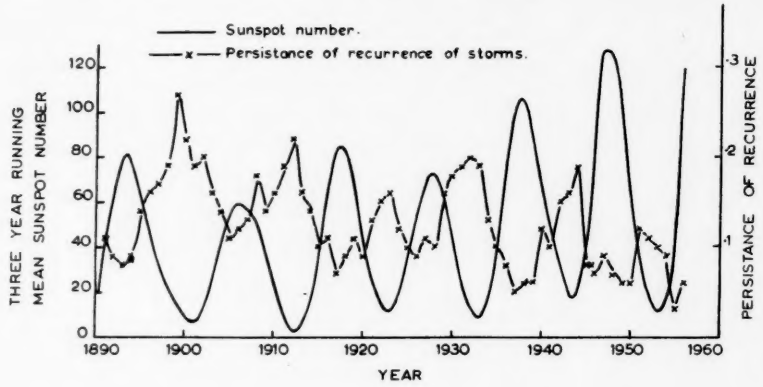


Fig. 2.

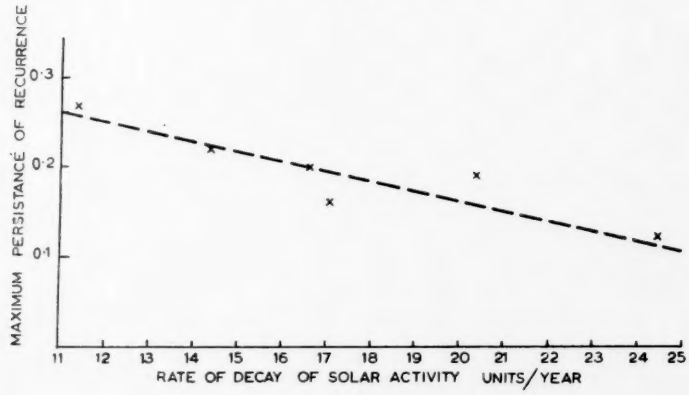


Fig. 3.

27 day pattern during the 1954 minimum period was the result of the phenomenally high rate of decay of the preceding cycle of solar activity.

CONCLUSION

By studying the recurrence of magnetic storms for the period 1890–1957 it is shown that the maximum value of persistence of recurrence of magnetic storms during the sunspot minimum period varies in an inverse manner with the rate of decay of solar activity to that minimum. It can be inferred that the lives of M-regions vary inversely as the rate of change of solar activity and that the

VOL. 17 1959/6

observed lack of success in the forecasting of radio propagational disturbances during the 1954 sunspot minimum period is a consequence of the high rate of decay of the present sunspot cycle.

Acknowledgements—The writer expresses his appreciation of the help he has received from his colleagues at Dominion Physical Laboratory, in particular Messrs. W. H. Ward and N. F. Barber.

References

Allen C. W.	1944	*Mon. Not. R. Astr. Soc.* **104,** No. 1, 13.
Babcock H. W. and Babcock H. D.	1955	*Nature, Lond.* **175,** 296.
Bartels J.	1932	*Terr. Magn. Atmos. Elect.* **37,** 1.
Chapman S. and Ferraro V. C. A.	1931–1933	*Terr. Magn. Atmos. Elect.* **36,** 77, 171; **37,** 147; **38,** 79.
Chapman S. and Bartels J.	1940	*Geomagnetism* Vol. II, Table A, p. T1. Clarendon Press, Oxford.
Mood A. M.	1940	*Ann. Math. Stat.* **11,** 367.
Waldmeier M.	1946	*Terr. Magn. Atmos. Elect.* **51,** 537.

OL.
17
9/60

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 320 to 324. Pergamon Press Ltd. Printed in Northern Ireland

A giant geomagnetic pulsation

J. VELDKAMP
Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(*Received* 3 *August* 1959)

Abstract—A giant geomagnetic pulsation, which occurred on 17 July 1958, was recorded by many European observatories between 50° and 60° geomagnetic latitude. The movements (with periods of about 100 sec) are coherent in only a small part of the area in which they are recorded. The greatest amplitudes (25 γ) are found in a strip following the circles of geomagnetic latitude. Comparisons are made with theories of OBAYASHI and JACOBS, and SCHOLTE.

ON 17 July 1958, a remarkable vibration of the geomagnetic field was recorded at many magnetic stations in Western Europe. The vibration was almost harmonic, with a period of somewhat more than 100 sec; it lasted for more than 1 hr, at some stations even for more than 3 hr. Although pulsations of the geomagnetic field are very common phenomena, vibrations with periods and amplitudes as large as those recorded on 17 July 1958 are rare. These vibrations have been described by ROLF (1931), by HARANG (1932, 1936, 1939) and by SUCKSDORFF (1939) as giant pulsations (g.p.). HARANG found that during the giant pulsation the perturbing vector moves in an elliptical orbit. The g.p. are confined to the auroral zones and can after him be ascribed either to periodic vertical movements in the ionosphere, or to a periodic intrusion of an ionizing agency in the ionosphere. KATO and WATANABE (1956) suggest that for a certain station a nearly harmonic spectrum of periods is observed. WHITHAM and LOOMER (1958) found that the g.p.s appear only to an extent considerably less than 2000 to 3000 km. They think that magnetohydrodynamic waves in the upper part of the ionosphere can best explain the regular oscillations, as was suggested by LEHNERT (1956).

The g.p. of 17th July 1958 offered an excellent opportunity for studying the properties of this disturbance, as the number of magnetic observatories which recorded this phenomenon was exceptionally large.

In Fig. 1 part of the quick-run magnetogram from the Netherlands Magnetic Observatory Witteveen is reproduced. The g.p. started at 0915 hours and ended at 1245 hours; the greatest movements occurred between 0915 hours and 1045 hours. The perturbing vector moved anti-clockwise in an elliptical horizontal path. The maximum amplitude was about 12 γ. The period was about 95 sec in the beginning; after 0930 hours the period increased to 105 sec and remained constant during a remarkably long time. The greatest axis of the ellipse was in the direction WNW to ESE; the ratio between the two axes was about 1 : 1·2. When comparing the geomagnetic pulsations H with the corresponding geoelectric pulsations E recorded at the same station Witteveen, we find a phase difference of 17° between corresponding maxima and minima, and an amplitude ratio E/H of 0·4 mV/γ km. These values fit rather well to the E/H curve valid for Witteveen (SCHOLTE and VELDKAMP, 1955).

Fig. 2 shows the geographical distribution of the geomagnetic stations in which

VOL.
17
1959/6

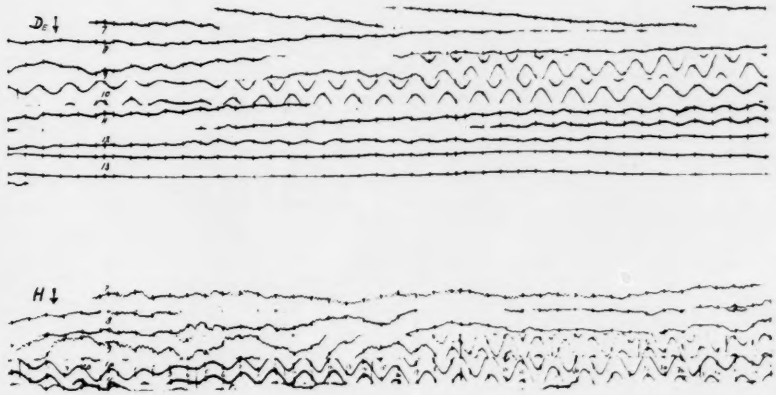


Fig. 1. Part of the La Cour quick-run magnetogram of 17 July 1958, recorded at the magnetic observatory Witteveen (*above*, declination, *below*, horizontal intensity). The dots in the record are minute-marks. The maximum double amplitude of the pulsation is 12 γ for D and 10 γ for H.

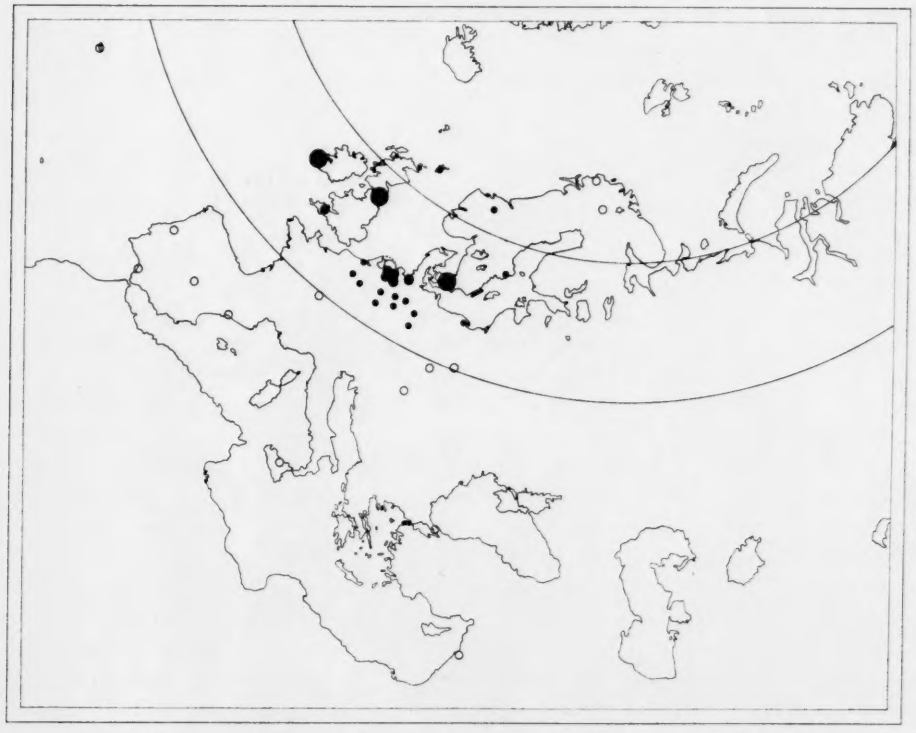


Fig. 2. Stations in which the giant pulsation of 17 July 1958, was recorded (black dots), and stations in which no g.p. was recorded (open circles). The size of the dots is a measure for the amplitude of the g.p. The great circles are drawn at geomagnetic latitudes of 50° and 60°. For the names of the stations see Table 1.

Table 1

Station	Co-ordinates		Time U.T.	Amplitude		
				H	*D*	*Z*
Dombâs	62°04′	9°07′	0915–1000	1γ	1γ	0γ
Lerwick*	60°08′	358°49′	0910–1020	2γ	6γ	0γ
Lovö*	59°21′	17°50′	0910–1145	8γ	6γ	6γ
Rude Skov	55°51′	12°27′	0915–1220	10γ	25γ	5γ
Eskdalemuir*	55°19′	356°48′	0915–1230	20γ	25γ	6γ
Helu	54°36′	18°48′	0915–1045	3γ	2γ	
Wingst*	53°45′	9°04′	0915–1230	12γ	13γ	0γ
Lathen	52°52′	7°19′	0915–1045	9γ	9γ	2γ
Witteveen*	52°49′	6°40′	0915–1245	10γ	12γ	0γ
Vorwald	52°37′	6°47′	0915–1045	5γ	8γ	0γ
Westerholte	52°33′	7°50′	0915–1045	5γ	4γ	1γ
Fallersleben	52°25′	10°44′	0915–1045	3γ	4γ	3γ
Elze	52°07′	9°45′	0915–1045	3γ	4γ	0γ
Niemegk*	52°04′	12°40′	0920–1040	2γ	2γ	1γ
Valentia*	51°56′	349°45′	0910–1220	10	20γ	25γ
Göttingen	51°33′	9°58′	0915–1045	3γ	2γ	1γ
Collmberg*	51°19′	13°00′	0930–1030	1γ	1γ	0γ
Schalksmühle	51°14′	7°32′	0915–1045	1γ	1γ	0γ
Hartland*	51°00′	4°30′	0900–1150	9γ	7γ	6γ
Ehringhausen	50°36′	8°24′	0915–1045	1γ	1γ	0γ
Manhay*	50°18′	5°41′	0935–1040	1γ	1γ	0γ
Dourbes*	50°06′	4°23′	0915–1050	1γ	2γ	0γ
Chambon la Forêt*	48°01′	2°16′	0930–1045	1γ		
Fredericksburg*	38°12′	282°38′	0945–1140	1γ	2γ	1γ

the g.p. of 17 July 1958 was recorded. They are all situated in Europe, except for one station in America (Fredericksburg). The dots represent stations in which the pulsation was recorded; open circles are stations in which no g.p. was registered. The greatest amplitudes (about 25 γ) were measured in the stations Valentia (Ireland), Eskdalemuir (Scotland) and Rude Skov (Denmark). Roughly speaking, the pulsation occurred in a strip which follows the circles of geomagnetic latitude.

Table 1 gives a summary of the maximum (double) amplitudes recorded; stations with a quick-run recorder have been marked by an asterisk. The time indicated is the total time during which the g.p. was recorded.

The giant pulsation occurred in the beginning of a period of moderate disturbance; the K_p figures for the first half of the day were 4_0, $5-$, 4_0 and $3+$. No real magnetic storm developed although the disturbed period was the recurrence of previous periods of storminess in the months of May and June. In the beginning of the day almost continously lasting pulsations (p.c.) with periods of about 12 sec and amplitudes of a few γ were recorded. They disappeared at 0930 hours with a sudden impulse (s.i.); after that the great periods and amplitudes of the g.p. began.

When comparing the records of different stations, a remarkable phenomenon is observed. Whereas the maxima and minima of the pulsations occur at practically

VOL. 17 1959/6

the same time for near stations, the correlation is completely lost for stations with distances greater than 500 km. The period of the vibrations is the same for all stations, but phase differences between the movements at different stations are found which vary over a large angle during the pulsation. The times of beginning and ending of the g.p. are also different for various stations. The ratio of the axes of the elliptical movement varies between 1·0 and 2·0 from one station to another. For most stations the movement of the perturbing vector is anti-clockwise, except for Lovö (Sweden) and Fredericksburg (U.S.A.), where the movement is clockwise. It is obvious that the direct cause of the g.p. cannot be the same in stations far apart.

COULOMB (1957) has tried to explain the pulsations by ascribing them to an electric dipole (or a vertical current) in the polar region. This would imply east–west geomagnetic variations. However, the movements observed during the g.p. of 17 July 1958 were elliptical and for some stations even almost circular. Moreover the movements were not simultaneous in stations far apart.

After the theory of OBAYASHI and JACOBS (1958) the geomagnetic pulsations are generally caused by hydromagnetic (or magneto-ionic) oscillations in the exosphere. The g.p. might indeed be ascribed to magneto-ionic oscillations along the lines of force of the geomagnetic field. The period observed (105 sec) fits very well to the period (100 sec) found by these authors for a geomagnetic latitude of 57°. The phenomenon of the whistling atmospherics has proved (STOREY, 1953) that the so-called whistlers are generally propagated through tubes of somewhat enhanced ionization; hereby the whistler can only be recorded in a limited area. The fact that the giant pulsation is coherent over only a small part of the area in which it was recorded, makes it difficult to suppose a tube or sheet of enhanced ionization in which the whole movement of the g.p. takes place.

SCHOLTE (1960) has shown that a disturbance in an ionized medium can be propagated through a homogeneous magnetic field along the lines of force without loss of energy in directions perpendicular to the field. We may suppose that this also holds for the dipole field of the earth. This means that a disturbance somewhere in the exosphere caused by an outburst on the sun, will be propagated along the lines of force while retaining its original dimensions.

So we might conclude that the area over which the g.p. of 17 July 1958 was recorded is an indication of the size of the area in which the exosphere was disturbed at the generation of the g.p. The loss of coherence of the vibrations over a rather small distance, and the shift of starting and ending times might be caused by the occurrences in the source of the disturbance; probably not the whole area was disturbed at the same moment.

The proof of the supposition of longitudinal propagation of the magneto-ionic disturbance might be delivered by the records of stations situated at the southern end of the lines of force. However, records of such stations have not come to my attention.

Acknowledgement—Magnetograms were thankfully received from the following observatories: Tromsö, Dombås, Lerwick, Eskdalemuir, Lovö, Abisko, Kiruna, Rude Skov, Hel, Belsk, Raciborz, Wingst, Vorwald, Schalksmühle, Göttingen,

Lathen, Westerholte, Fallersleben, Elze, Ehringhausen, Witteveen, Niemegk, Valentia, Collmberg, Hartland, Manhay, Dourbes, Chambon-la-Forêt, Churchill, College, Baker Lake, Resolute Bay, Meanook, Agincourt, Yellow Knife, Sitka, Victoria, Fredericksburg, Ebro, Toledo, Coimbra, San Miguel, Ksara, Istanbul Kandilli, San Fernando, M'Bour, Bangui, Tamanrasset, Luanda, Elisabethville, Tucson, San Juan, Honolulu, Tananarive, Manila, Quetta, Kuyper, Huancayo, Hermanus, Pilar, Watheroo.

REFERENCES

COULOMB J.	1957	*Ann. Géophys.* **13,** 91.
HARANG L.	1932	*Terr. Magn. Atmos. Elect.* **37,** 57.
HARANG L.	1936	*Terr. Magn. Atmos. Elect.* **41,** 329.
HARANG L.	1939	*Terr. Magn. Atmos. Elect.* **44,** 17.
KATO Y. and WATANABE T.	1956	*Sci. Rep. Tokohu Univ. Ser.* 5, *Geoph.* **8, 1.**
LEHNERT B.	1956	*Tellus* **8,** 241.
OBAYASHI T. and JACOBS J. A.	1958	*Geophys. J.* **1,** 53.
OBAYASHI T.	1958	*Ann. Géophys.* **14,** 464.
ROLF B.	1931	*Terr. Magn. Atmos. Elect.* **36,** 9.
SCHOLTE J. G. and VELDKAMP J.	1955	*J. Atmosph. Terr. Phys.* **6,** 33.
SCHOLTE J. G.	1960	*J. Atmosph. Terr. Phys.* In press.
STOREY L. R. O.	1953	*Phil. Trans. Roy. Soc.* A **246,** 113.
SUCKSDORFF E.	1939	*Terr. Mag. and Atm. Elec.* **44,** 337.
WHITHAM K. and LOOMER E. I.	1958	*Publ. Dominion Obs.* **19,** 265.

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 325 to 336. Pergamon Press Ltd. Printed in Northern Ireland

On the theory of giant pulsations

J. G. J. SCHOLTE
Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(*Received* 3 *August* 1959)

Abstract—The propagation of a magneto-ionic disturbance originating in a small region of the exosphere has been investigated; it appears that only a rotational movement perpendicular to the geomagnetic field is propagated without any great loss of energy. The amplitude of a compressional movement perpendicular to this field decreases with increasing distance to the source, whereas a movement parallel to the magnetic field is propagated as a sound-wave without any magnetic effect. Therefore the observed giant pulsations, which are ascribed to standing waves along the lines of force, are mainly caused by the rotational part of the primary disturbance and the properties of these variations are closely connected with the properties of that part of the original movement.

INTRODUCTION

IN order to explain the occurrence of geomagnetic oscillations with periods up to 100 sec, a theory has been put forward which may be briefly described as follows: the primary cause of these oscillations is a disturbance of the ionized medium (the exosphere) surrounding the earth as a distance of several earth's radii; this disturbance may probably be brought about by a cloud of ionized particles erupted by the sun. This perturbation and the associated disturbance of the geomagnetic field are mainly propagated in the direction of the magnetic field with the velocity of Alfvén-waves. Arriving at the surface of the earth this magneto-ionic perturbation will be totally reflected, so that the variation of the geomagnetic field will travel to and fro along the magnetic lines of force between their terminals at the earth's surface.

Hence a magnetogram shows a periodical variation with a period equal to the time necessary to travel twice along a geomagnetic line of force with the Alfvén-velocity. It follows immediately that this period increases with increasing geomagnetic latitude, which circumstance has been confirmed by an investigation of many giant pulsations by OBAYASHI and JACOBS (1958). Again the travel time along a line of force depends on the Alfvén-velocity which is a function of the ionic density; therefore the observed periods give us some information about this density, analogous to the information contained in the seismological travel–time tables. The above-mentioned authors derive from the periods of the giant pulsations a picture of the ionic density of the exosphere as a function of the height.

Apart from the period a vibration is also characterized by its amplitude and if vibrations at different points are compared with each other, by the phase-differences. In a recent study VELDKAMP (1960) examined this aspect of the phenomenon by investigating the direction, amplitude and phase of the magnetic variation observed at different stations during the occurrence of a giant pulsation on 17 July 1958.

GENERAL EQUATIONS

In connexion with this investigation it is of interest to extend the above theory somewhat in order to obtain some results relevant to these properties. The problem is rather simply stated in the following way: suppose that an ionized medium

existing in a magnetic field is disturbed at one point; the question is to calculate the magnetic field in any point of this medium.

The ionized medium is defined by the density ρ, the magnetic permeability μ, the dielectric constant ε, the conductivity σ and the coefficient of viscosity ν. We suppose that this medium extends to infinity in all directions and that the magnetic field B_0 is homogeneous if the medium is at rest. The equation of motion of this medium is

$$\rho \frac{Dv}{Dt} = -\nabla p + \rho \nu \Delta v + i \wedge B$$

where v = the velocity, p = the pressure, B_1 = the primary field B_0 + induced field.

Taking the z-axis in the direction of B_0 and neglecting all quadratic terms (products of two velocity-components or products of i and the induced field) we obtain

$$\rho \frac{\partial v_x}{\partial t} + \frac{\partial p}{\partial x} - \rho \nu \Delta v_x = B_0 i_y$$

$$\rho \frac{\partial v_y}{\partial t} + \frac{\partial p}{\partial y} - \rho \nu \Delta v_y = -B_0 i_x$$

$$\rho \frac{\partial v_z}{\partial t} + \frac{\partial p}{\partial z} - \rho \nu \Delta v_z = 0.$$

In accordance with the nature of the problem we introduce the quantities $\operatorname{curl}_z v$ and $\operatorname{div} v$, thereby emphasizing the importance of the z-direction; the equations change into

$$\rho\left(\frac{\partial}{\partial t} - \nu\Delta\right) \operatorname{curl}_z v = -B_0\left(\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y}\right)$$

$$\rho\left(\frac{\partial}{\partial t} - \nu\Delta\right) \operatorname{div} v = B_0 \operatorname{curl}_z i - \Delta p$$

$$\rho\left(\frac{\partial}{\partial t} - \nu\Delta\right) v_z = -\frac{\partial p}{\partial z}$$

Assuming the exosphere to be isothermal we have $\nabla p = (\partial p/\partial \rho)\nabla\rho$ or as $\partial p/\partial\rho$ = the square of the velocity v_s of sound: $\nabla p = v_s^2 \nabla\rho$. The equation of continuity $\rho \operatorname{div} v + \partial\rho/\partial t = 0$ then becomes $\partial p/\partial t = -\rho v_s^2 \operatorname{div} v$.

By means of this relation p can be eliminated from the equations, and we obtain three equations connecting the velocity v with the current density i:

$$\rho\left(\frac{\partial}{\partial t} - \nu\Delta\right) \operatorname{curl}_z v = -B_0\left(\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y}\right)$$

$$\rho\left(\frac{\partial^2}{\partial t^2} - v_s^2\Delta - \nu \frac{\partial}{\partial t}\Delta\right) \operatorname{div} v = B_0 \frac{\partial \operatorname{curl}_z i}{\partial t}$$

$$\left(\frac{\partial^2}{\partial t^2} - \nu \frac{\partial}{\partial t}\Delta\right) v_z - v_s^2 \frac{\partial \operatorname{div} v}{\partial z} = 0.$$

Apart from this hydrodynamical connexion between v and i these quantities are also connected by electromagnetic relations. These are expressed by equations containing v, i, B and the electric field E, from which by eliminating B and E we obtain a second relation between v and i.

The law of Ohm is expressed by: $i = \sigma(E + v \wedge B)$

or: $$i_x = \sigma(E_x + B_0 v_y),\ i_y = \sigma(E_y - B_0 v_x),\ i_z = \sigma E_z$$

hence $$\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y} = \sigma\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\right) + \sigma B_0 \operatorname{curl}_z v, \tag{1}$$

and $$\operatorname{curl}_z i = \sigma \operatorname{curl}_z E - \sigma B_0\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right). \tag{2}$$

Again i and E satisfy Maxwell's equation $i + \varepsilon(\partial E/\partial t) = \operatorname{curl} H$ and finally H and E satisfy the equation of induction $\mu(\partial H/\partial t) = -\operatorname{curl} E$. Elimination of H yields two equations:

$$\operatorname{div} i = -\varepsilon \frac{\partial \operatorname{div} E}{\partial t}$$

and $$-\operatorname{curl}\operatorname{curl} E = \mu \frac{\partial}{\partial t}\left(i + \varepsilon \frac{\partial E}{\partial t}\right)$$

or $$\mu \frac{\partial i}{\partial t} = -\nabla \operatorname{div} E + \beta E$$

with $$\beta = \Delta - \varepsilon\mu\, \partial^2/\partial t^2.$$

Using these relations the quantities $\operatorname{curl}_z i$ and $\partial i_x/\partial x + \partial i_y/\partial y$, occurring in the equations of motion, are readily expressed, respectively, in $\partial v_x/\partial x + \partial v_y/\partial y$ and $\operatorname{curl}_z v$.

1. As $\operatorname{div} i = -\varepsilon \dfrac{\partial \operatorname{div} E}{\partial t}$ and $i_z = \sigma E_z$ we have

$$\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y} = -\varepsilon \frac{\partial}{\partial t}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\right) - \left(\sigma + \varepsilon \frac{\partial}{\partial t}\right)\frac{\partial E_z}{\partial z}.$$

Again as $\mu \dfrac{\partial i_z}{\partial t} = -\dfrac{\partial \operatorname{div} E}{\partial z} + \beta E_z$ and $i_z = \sigma E_z$ it follows

$$\frac{\partial}{\partial z}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\right) = \left(\beta - \frac{\partial^2}{\partial z^2} - \mu\sigma \frac{\partial}{\partial t}\right) E_z.$$

Elimination of E_z yields

$$\left(\beta - \frac{\partial^2}{\partial z^2} - \mu\sigma \frac{\partial}{\partial t}\right)\left(\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y}\right) = -\left\{\varepsilon \frac{\partial}{\partial t}\left(\beta - \mu\sigma \frac{\partial}{\partial t}\right) + \sigma \frac{\partial^2}{\partial z^2}\right\}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\right);$$

using (1) we obtain:

$$\alpha\left(\beta - \sigma\mu\,\frac{\partial}{\partial t}\right)\left(\frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y}\right) = \sigma B_0\left\{\alpha\beta - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\right\}\operatorname{curl}_z v \qquad (3)$$

where $\alpha = 1 + \dfrac{\varepsilon}{\sigma}\dfrac{\partial}{\partial t}$.

2. From $\mu\,\dfrac{\partial i}{\partial t} = -\nabla \operatorname{div} + \beta E$ we derive $\mu(\partial \operatorname{curl}_z/\partial t) = \beta \operatorname{curl}_z$; by means of (2) this becomes

$$\left(\beta - \sigma\mu\,\frac{\partial}{\partial t}\right)\operatorname{curl}_z i = -\sigma\beta B_0\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right). \qquad (4)$$

Using (3) and (4) we eliminate i from the equations of motion, which then become:

$$\left[\alpha\left\{\beta\left(1 + \frac{\rho}{\sigma {B_0}^2}\,\frac{\partial}{\partial t} - \frac{\rho\nu}{\sigma {B_0}^2}\,\Delta\right) - \frac{\mu\rho}{{B_0}^2}\,\frac{\partial}{\partial t}\left(\frac{\partial}{\partial t} - \nu\Delta\right)\right\} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\right]\operatorname{curl}_z v = 0 \quad \text{(I)}$$

$$\left\{\left(\frac{\rho\beta}{\sigma {B_0}^2} - \frac{\rho\mu}{{B_0}^2}\,\frac{\partial}{\partial t}\right)\left(\frac{\partial^2}{\partial t^2} - {v_s}^2\Delta - \nu\,\frac{\partial}{\partial t}\,\Delta\right) + \beta\,\frac{\partial}{\partial t}\right\}\operatorname{div} v - \beta\,\frac{\partial^2 v_z}{\partial t \partial z} = 0 \quad \text{(II)}$$

$$\left(\frac{\partial^2}{\partial t^2} - \nu\,\frac{\partial}{\partial t}\,\Delta\right) v_z - {v_s}^2\,\frac{\partial \operatorname{div} v}{\partial z} = 0 \qquad \text{(III)}$$

A similar deriviation has been given by BAÑOS (1955).

In these equations the terms with the factors $1/\sigma$ or ν represent energy-dissipating effects; denoting these terms by $\{1/\sigma, \nu\}$, the equations may be written in more surveyable forms:

$$\left\{\rho\,\frac{\partial^2}{\partial t^2} - {B_0}^2\left(\frac{1}{\mu}\,\frac{\partial^2}{\partial z^2} - \varepsilon\,\frac{\partial^2}{\partial t^2}\right)\right\}\operatorname{curl}_z v + \{1/\sigma, \nu\} = 0$$

$$\frac{\partial}{\partial t}\left\{\left(\rho\,\frac{\partial^2}{\partial t^2} - \rho {v_s}^2\Delta\right)\operatorname{div} v - {B_0}^2\left(\frac{1}{\mu}\,\Delta - \varepsilon\,\frac{\partial^2}{\partial t^2}\right)\left(\operatorname{div} v - \frac{\partial v_z}{\partial z}\right)\right\} + \{1/\sigma, \nu\} = 0$$

$$\left(\rho\,\frac{\partial^2 v_z}{\partial t^2} - \rho {v_s}^2\,\frac{\partial \operatorname{div} v}{\partial z}\right) + \{\nu\} = 0$$

In the first equation, only terms relating to the inertial and the electromagnetic force occur, whereas in the second equation a term connected with the pressure appears.

It is evident that a solution v of these equations consists of two parts: one part is a purely horizontal movement ($v_z = 0$) which, according to (III), is without divergence; of the other part the rotation is purely horizontal ($\operatorname{curl}_z v = 0$).

The movement v is determined by these equations and by the initial or boundary conditions; in this problem the movement is prescribed in a limited region of the medium. It is difficult to specify the form of this primary movement as it depends

VOL. 17 1959/6

on the way in which it is caused; however, as any movement v is defined by the three quantities v_z, $\mathrm{curl}_z\, v$ and $\mathrm{div}\, v$ the primary movement can in any case be decomposed into a horizontal component without divergence ($v_z = 0$, $\mathrm{div}\, v = 0$) and a component with a horizontal rotation ($\mathrm{curl}_z\, v = 0$). Each of these partial disturbances is propagated in the medium in a different way: $\mathrm{curl}_z\, v$ satisfying (I) and $\mathrm{div}\, v$ and v_z satisfying (II) and (III).

Consequently the magnetic variation $B - B_0$ resulting from the disturbance is also brought about in two different ways; analogous to the situation in seismology there is here no simple connexion between the primary disturbance and the observed phenomenon, as there are two different velocities of propagation.

This result obtains for any homogeneous and isotropic ionized medium which is moving in a homogeneous magnetic field.

In the following we will consider the two kinds of propagation separately, neglecting the energy-dissipating factors $1/\sigma$ and ν.

1. Rotational Movement ($\mathrm{curl}_z\, v$)

We first consider the movement perpendicular to the primary field B_0; as $v_z = 0$ and $\mathrm{div}\, v = 0$ we have

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0$$

or

$$v_x = -\frac{\partial \psi}{\partial y}, \quad v_y = +\frac{\partial \psi}{\partial x},$$

where ψ is a function of x, y, z, t. The quantity $\mathrm{curl}_z\, v$ becomes $(\partial^2/\partial x^2 + \partial^2/\partial y^2)\psi$ from which follows that ψ must satisfy equation (I): hence, in this approximation ($1/\sigma$ and $\nu = 0$):

$$\left\{\frac{\partial^2}{\partial z^2} - \left(\varepsilon\mu + \frac{\rho\mu}{{B_0}^2}\right)\frac{\partial^2}{\partial t^2}\right\}\psi = 0.$$

Introducing the velocity v_e of electromagnetic waves $[= 1/\sqrt{(\varepsilon\mu)}]$ and the velocity v_a of Alfvén-waves $[= B_0/\sqrt{(\rho\mu)}]$ the equation becomes

$$\left\{\frac{\partial^2}{\partial z^2} - \left(\frac{1}{{v_e}^2} + \frac{1}{{v_a}^2}\right)\frac{\partial^2}{\partial t^2}\right\}\psi = 0$$

with the general solution: $\psi = F(t \pm z/w,\ x,\ y)$ with $1/w^2 = 1/{v_a}^2 + 1/{v_e}^2$. The propagation takes place in the direction of the magnetic field, which is a transverse propagation as the motion is perpendicular to this direction. The associated electric field is determined by $E + v \wedge B = 0$, which gives $E_x = -B_0 \partial F/\partial x$, $E_y = -B_0 \partial F/\partial y$, $E_z = 0$, and the induced magnetic field by

$$H_x = \mp\frac{B_0}{\mu w}\frac{\partial F}{\partial y}, \qquad H_y = \pm\frac{B_0}{\mu w}\frac{\partial F}{\partial x}, \qquad H_z = 0.$$

Apparently the electric field is perpendicular to v, and the induced magnetic field parallel to the velocity. Neglecting $1/v_e$ with respect to $1/v_a$ the induced field becomes simply $H = \pm v\sqrt{(\rho/\mu)}$.

If we assume for $z > 0$ $F(t - z/v_a, x, y)$ to be a function which is identically zero if $t < z/v_a$ and not identically zero if $t \geqslant z/v_a$ then the movement v determined by this function is propagated in the direction of the primary field with the velocity v_a. The accompanying electromagnetic field is proportional to the movement v; in particular the extension of the primary disturbance in the directions perpendicular to the primary field is identical to the extension in these directions of the e.m. field.

Concluding this part of this investigation we arrive at the following conclusions:

(1) As far as the rotational, transverse disturbance is concerned the propagation of such a perturbation, accompanied by an e.m. field, proceeds along the lines of force without any loss of energy apart from the effects of viscosity and electrical resistance.

(2) Although the geomagnetic field is not homogeneous it may be expected that the inhomogeneity does not greatly affect the above results. Then we can state that the region where the original disturbance of the exosphere takes place is projected along the lines of force; accordingly inside the area of the earth's surface bounded by these lines, magnetic variations will occur.

2. MOVEMENT WITH $\operatorname{curl}_z v = 0$

The second part of the movement is characterized by $\operatorname{curl}_z v = 0$; the motion must satisfy equations (II) and (III) which become, if we neglect the terms with $1/\sigma$ and ν:

$$\left\{\frac{1}{{v_a}^2}\left(\frac{\partial^2}{\partial t^2} - {v_s}^2\Delta\right) + \left(\Delta - \frac{1}{{v_c}^2}\frac{\partial^2}{\partial t^2}\right)\right\} \operatorname{div} v + \left(\Delta - \frac{1}{{v_c}^2}\frac{\partial^2}{\partial t^2}\right)\frac{\partial v_z}{\partial z} = 0$$

$$\frac{1}{{v_s}^2}\frac{\partial^2 v_z}{\partial t^2} = \frac{\partial \operatorname{div} v}{\partial z}.$$

By elimination of either v_z or div v the first equation changes into

$$\left\{\frac{1}{{v_a}^2}\frac{\partial^2}{\partial t^2}\left(\frac{1}{{v_s}^2}\frac{\partial^2}{\partial t^2} - \Delta\right) + \left(\frac{1}{{v_c}^2}\frac{\partial^2}{\partial t^2} - \Delta\right)\left(\frac{1}{{v_s}^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2}\right)\right\} \begin{cases}\operatorname{div} v \\ v_z\end{cases} = 0. \quad \text{(IIa)}$$

A motion with $\operatorname{curl}_z v = 0$ is generally represented by the following expressions:

$$v_x = \frac{\partial \phi}{\partial x} - \frac{\partial^2 \chi}{\partial x \partial z}, \qquad v_y = \frac{\partial \phi}{\partial y} - \frac{\partial^2 \chi}{\partial y \partial z}, \qquad v^2 = \frac{\partial \phi}{\partial z} + \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\chi$$

where ϕ and χ are functions of x, y, z, t.

As div v $(= \Delta\phi)$ and v_z both satisfy equation (IIa) the functions ϕ and χ must also satisfy this equation. Again these two functions are related by equation (III), which becomes

$$\frac{\partial}{\partial z}\left(\frac{1}{{v_s}^2}\frac{\partial^2}{\partial t^2} - \Delta\right)\phi + \frac{1}{{v_s}^2}\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\chi = 0. \quad \text{(IIIa)}$$

Contrary to equation (1) in both equations (IIa) and (IIIa) the operator $\partial^2/\partial x^2 + \partial^2/\partial y^2$ appears; therefore the propagation of this movement depends also on the properties of this movement in directions perpendicular to the magnetic lines of force.

VOL. 17 1959/6

We now consider a simple solution ϕ of (IIa) namely one with the properties $\partial^2\phi/\partial t^2 = -\omega^2\phi$, $(\partial^2/\partial x^2 + \partial^2/\partial y^2)\phi = -k_h^2\phi$ and $\partial^2\phi/\partial z^2 = -k_v^2\phi$, where ω, k_h and k_v are constants. The solution is $\phi = \exp(i\omega t + ik_v z) f(x, y, k_h)$ where f satisfies

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2 + k_h^2) f = 0 \tag{5}$$

Denoting $\omega^2/v_a^2, \omega^2/v_c^2, \omega^2/v_s^2$, respectively, by k_a^2, k_c^2, k_s^2 equation (IIa) becomes:

$$k_a^2(k_s^2 - k_h^2 - k_v^2) + (k_c^2 - k_h^2 - k_v^2)(k_s^2 - k_v^2) = 0 \tag{6}$$

The connexion (6) between the two parameters k_h^2 and k_v^2 is graphically

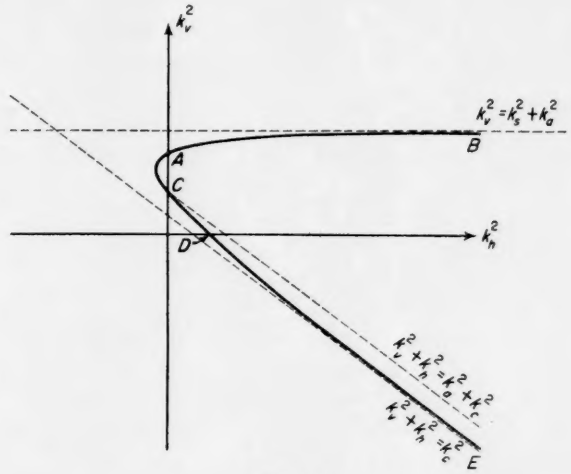


Fig. 1.

represented by a hyperbola (Fig. 1). For positive values of k_h^2 this curve consists of two parts: AB and CE; on the first part k_v^2 starts at $k_h^2 = 0$ with the value k_s^2 and approaches with increasing k_h^2 the asymptotic value $k_s^2 + k_a^2$. Hence on AB we have

$$k_s^2 + k_a^2 > k_v^2 \geqslant k_s^2 \qquad \text{or} \qquad k_s^2(1 + v_s^2/v_a^2) > k_v^2 \geqslant k_s^2$$

as in actual conditions the velocity of sound is much smaller than the velocity of Alfvén-waves the value of k_v^2 is on this part of the curve virtually equal to k_s^2. On the other part k_v^2 begins (in C) with the value $k_a^2 + k_c^2$, is equal to zero when

$$k_h^2 = \frac{k_s^2(k_a^2 + k_c^2)}{k_s^2 + k_a^2}$$

and approaches with increasing k_h^2 the value $k_c^2 - k_h^2$.

On the traject CD the inequality

$$k_a^2 + k_c^2 \geqslant k_v^2 + k_h^2 > (k_a^2 + k_c^2)/(1 + v_s^2/v_a^2)$$

obtains; using again $v_s/v_a \ll 1$ we conclude that on this traject $k_v^2 + k_h^2$ is approximately constant. As the velocity v_c of electromagnetic waves is of course

much greater than v_a we must in this approximation neglect k_c^2 too; therefore $k_r^2 + k_h^2 \simeq k_a^2$ on the traject CD.

On the remaining part DE the value of $k_v^2 + k_h^2$ lies between k_a^2 and k_c^2, and while $k_h^2 > k_a^2$, it can be easily shown that the above approximation $k_v^2 + k_h^2 \simeq k_a^2$ remains valid if $k_h^2 \ll k_s^2$. However, if $k_h^2 \simeq k_s^2$ the value of k_v^2 is about equal to $-k_s^2$, so that the function ϕ decreases exponentially with z: $\phi \sim \exp(-k_s z)$. At a distance $z \gg 1/k_s$ the amplitude of this function is vanishingly small. Therefore if we use throughout the approximation $k_v^2 + k_h^2 \simeq k_a^2$ we obtain a value of ϕ which is approximately true, provided the distance z is great in comparison to the length of a sound-wave with the frequency in question. With a value $v_s \simeq 1$ km/sec the distance z, measured in units of 100 km, must satisfy $z \gg T/625$ where T is the period in seconds; this condition is in the phenomenon under discussion always fulfilled, the height of the exosphere being at least several hundreds of kilometres.

As, for each positive value of k_h^2, two values of k_v^2 exist, namely one which is about equal to k_s^2 and one which satisfies approximately the equation $k_v^2 + k_h^2 = k_a^2$, we have also two solutions $\phi = \exp(-i\omega t + ik_v z) f(x, y, k_h)$. Discussing these two waves separately we start with the first case: the function $f(x, y, k_h)$ must fulfill equation (5); as k_v is constantly equal to k_s the value of k_h is quite arbitrary, so that $f(x, y)$ may be any linear combination of functions satisfying (5), for arbitrary values of k_h. In other words $f(x, y)$ is an arbitrary function.

The movement v is determined by ϕ and χ and the function χ is obtained by solving (IIIa), which yields quite easily $\chi = \phi/ik_s$. It follows that $\phi - \partial\chi/\partial z = 0$ and then $v_x = v_y = 0$; therefore the movement is in the direction of the magnetic field and the propagation is longitudinal. Again the electric field vanishes (as $E + v \wedge B_0 = 0$) and consequently the variable magnetic field is also equal to zero.

The result is quite simply a sound wave which is longitudinally propagated along the lines of force without any magnetic disturbance.

Finally we consider the solution connected with the second root k_v^2; from equation (IIIa) we obtain

$$\chi = i \frac{k_v}{k_h^2} \left(1 - \frac{k_a^2 + k_c^2}{k_s^2}\right) \phi$$

or approximately

$$\chi \simeq i \frac{k_v}{k_h^2} \phi.$$

The components of the velocity appear to be

$$v_x \simeq \frac{k_a^2}{k_h^2} \frac{\partial\phi}{\partial x}, \qquad v_y \simeq \frac{k_a^2}{k_h^2} \frac{\partial\phi}{\partial y}, \qquad v_z \simeq 0$$

hence in first approximation the movement is irrotational and perpendicular to the magnetic field.

In order to investigate the properties of this movement, the function $f(x, y, k_h)$ which satisfies equation (5) must be specified; supposing the movement to be symmetrical with respect to the z-axis this equation becomes

$$(\partial^2/\partial r^2 + 1/r\ \partial/\partial r + k_h^2) f(r, k_h) = 0$$

VOL. 17 1959/6

with $r = (x^2 + y^2)^{\frac{1}{2}}$. Hence f is a Bessel-function $I_0(k_h r)$ of zero order and the wave function ϕ is $\phi = \exp(-i\omega t + ik_v z) I_0(k_h r)$ where k_h and k_v are connected by (6).

We now generalize this solution by superposition of various functions ϕ related to different values of k_h; in this way it is possible to construct a function ϕ which is equal to zero in all points of the plane $z = 0$ outside a circle with radius a and equal to some value $p \exp(-i\omega t)$ inside this circle. The function f is then expressed by the integral

$$f = pa \int_0^\infty I_0(k_h r) I_1(k_h a)\, dk_h .$$

The velocity v as given by $\partial\phi/\partial z$ is in this case equal to zero in every point of the plane $z = 0$ with the exception of points on the circle $r = a$ where the velocity is infinite.

Following the usual procedure to obtain a disturbance concentrated in a point we diminish the area of the circle, at the same time increasing the amplitude p in such a way that $\frac{1}{2}pa^2$ remains constant (q); we then obtain

$$f = q \int_0^\infty I_0(k_h^2) k_h\, dk_h .$$

Applying the operator $q \int_0^\infty k_h\, dk_h$ to the potential $\phi = \exp(-i\omega t + ik_v z) I_0(k_h r)$ this becomes

$$\phi = q \exp(-i\omega t) \int_0^\infty \exp(ik_v z) I_0(k_h r) k_h\, dk_h .$$

In this integral k_v^2 has the value of the ordinate of some point on the traject CE of the hyperbola (Fig. 1), and this value is approximately equal to $k_a^2 - k_h^2$ only if $k_h \ll k_s$. Nevertheless we take in the integral $k_v = \sqrt{(k_a^2 - k_h^2)}$ throughout the entire interval of integration thereby obtaining a result which is approximately valid if $k_s z \gg 1$.

According to a classical result the value of ϕ is

$$\phi = -\frac{\partial}{\partial z}\left(q \frac{\exp(-i\omega t + ik_a R)}{R}\right)$$

with $R = (x^2 + y^2 + z^2)^{\frac{1}{2}}$,

or
$$\phi = q \cos\theta \frac{1 - ik_a R}{R} \exp[-i\omega(t - R/v_a)]$$

where $\cos\theta = z/R$.

The velocity v is in the r-direction and (at distances $z \gg 1/k_s$) equal to

$$v_r = -qk_a^2 \exp(-i\omega t) \int_0^\infty \exp(ik_v z) I_1(k_h r)\, dk_h$$

or

$$v_r = -qk_a^2 \exp(-i\omega t) \frac{1}{r}[\exp(ik_a z) - \cos\theta \exp(ik_a R)].$$

The corresponding e.m. field is given by

$$E_x = -B_0 v_y, \quad E_y = +B_0 v_x, \quad E_z = 0$$

and $$H_x = \frac{iB_0}{\mu\omega} k_v v_x, \qquad H_y = \frac{iB_0}{\mu\omega} k_v v_y, \qquad H_z = -\frac{iB_0}{\mu\omega}\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right).$$

If the velocity-potential $\phi = \exp(-i\omega t + ik_v z) I_0(k_h r)$ these expressions lead to an electric field in the azimuthal (α)-direction equal to

$$E_\alpha = -B_0 \frac{k_a^2}{k_h^2} I_1(k_h r) \exp(-i\omega t)$$

and a magnetic field, given by

$$H_z = \frac{B_0 k_a^2}{\mu\omega} \frac{k_v}{k_h} I_1(k_h r) \exp(-i\omega t), \qquad H_\alpha = 0,$$

$$H_z = i \frac{B_0 k_a^2}{\mu\omega} I_0 k_h r \exp(-i\omega t).$$

Application of the operator $q \int_0^\infty k_h \, dk_h$ results in:

$$E_\alpha = \frac{qB_0 k_a^2}{r} [\exp(ik_a z) - \cos\theta \exp(ik_a R)] \exp(-i\omega t)$$

$$H_r = -i \frac{qB_0 k_a^2}{\mu\omega r} \frac{\partial}{\partial z} [\exp(ik_a z) - \cos\theta \exp(ik_a R)] \exp(-i\omega t),$$

$$H_z = -i \frac{qB_0 k_a^2}{\mu\omega} \frac{\partial}{\partial z}\left[\frac{\exp(ik_a R)}{R}\right] \exp(-i\omega t).$$

The components of H may also be written in the form

$$H_r = \frac{qB_0 k_a^2}{\mu\omega}\left[\frac{k_a}{r} \exp(ik_a z) - \frac{k_a z^2}{r R^2} \exp(ik_a R) + \frac{ir}{R^3} \exp(ik_a R)\right] \exp(-i\omega t)$$

$$H_z = \frac{qB_0 k_a^2}{\mu\omega}\left[k_a + \frac{k_a z}{R^2} \exp(ik_a R) + \frac{iz}{R^3} \exp(ik_a R)\right] \exp(-i\omega t).$$

This e.m. field consists of two parts:
(1) The field

$$E_\alpha = \frac{qB_0 k_a^2}{r} \exp[-i\omega(t - z/v_a)],$$

$$H_r = \frac{qB_0 k_a^2}{\mu v_a r} \exp[-i\omega(t - z/v_a)]$$

which is connected with the velocity

$$v_r = -\frac{qk_a^2}{r} \exp[-i\omega(t - z/v_a)].$$

It follows that $|H_r|/|E_\alpha| = \mu v_a$ and that the magnetic field is proportional to the velocity: $H_r/v_r = -\sqrt{(\rho/\mu)}$. The same relations exist in the case of the primary

VOL. 17 1959/6

disturbance $\text{curl}_z\, v$; the only difference between these two cases is the decrease of the quantities with $1/r$, which does not exist in the case of a disturbance $\text{curl}_z\, v$ concentrated at one point.

(2) The second part of the e.m. field is given by the remaining terms, which all contain the phase-factor $\exp[-i\omega(t - R/v_a)]$. As these terms obviously describe

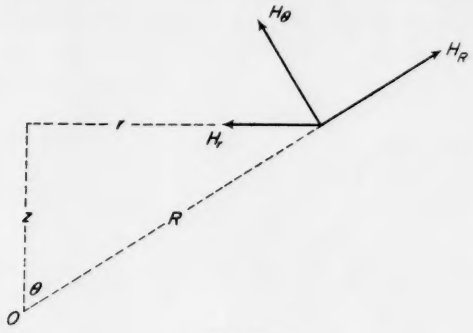


Fig. 2.

a wave which is radially propagated from the centre of disturbance it is convenient to use spherical co-ordinates (Fig. 2): R, θ, α

$$E_\alpha = -qB_0 {k_a}^2 \operatorname{cotg}\theta \frac{\exp[-i\omega(t - R/v_a)]}{R}$$

$$H_R = i\frac{qB_0 {k_a}^2}{\mu\omega}\frac{\exp[-i\omega(t - R/v_a)]}{R^2}, \qquad H_\theta = -\frac{qB_0 {k_a}^3}{\mu\omega}\operatorname{cotg}\theta\frac{\exp[-i\omega(t - R/v_a)]}{R}.$$

It is evident that this wave, travelling in any radial direction, is of no importance to the process of generating standing waves in the lines of force.

Conclusion

In the course of the above investigation (the original disturbance assumed to occur in the point $R = 0$) it appeared that we must consider three parts of the disturbance separatedly, namely v_z, div v and $\text{curl}_z\, v$. Each of these components is propagated in a different way:

(1) The movement v_z parallel to the magnetic field is simply propagated as a sound wave without any magnetic effect.

(2) The compressional movement div v becomes at some distance z ($\gg$ length of a sound wave) approximately perpendicular to the magnetic field; the motion and the accompanying magnetic disturbance are propagated with the Alfvén-velocity. This propagation proceeds partly in the direction of the field, the amplitude decreasing as $1/r$; while travelling to and fro along the lines of force the amplitude does not decrease, so that this disturbance will for small values of $1/r$ contribute to the generation of standing waves.

(3) The rotational movement curl_z is also perpendicular to the primary field and is propagated along the lines of force without any loss of energy (apart from

the effects of inhomogeneity of the primary field and of the viscosity and resistance of the ionized medium).

As has been mentioned at the beginning of this paper, VELDKAMP recently studied the geographical distribution of a giant pulsation; it appears that the correlation between the registrations at two stations diminishes with increasing distance and disappears altogether if the stations are wide apart. According to the above theory the original disturbance in the exosphere must show a similar pattern. As this disturbance is probably caused by a cloud of ionized particles this irregular pattern is understandable.

It is well known that these vibrations of the geomagnetic field are elliptically polarized (the ratio of the axes varying, according to VELDKAMP, from 1·0 to 2·0). As we have seen that these vibrations are mainly brought about by the component $\mathrm{curl}_z\, v$ this behaviour of the magnetic disturbance is not unexpected. Moreover it is often assumed that the ionized particles travel in spirals around the lines of force; thus the rotating character of the phenomenon may be connected with this circumstance.

REFERENCES

BAÑOS J. A.	1955	*Phys. Rev.* **97,** 1435.
OBAYASHI T. and JACOBS J. A.	1958	*Geophys. J.* **1,** 53.
VELDKAMP J.	1960	*J. Atmosph. Terr. Phys.*

VOL.
17
1959/6

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 337 to 344. Pergamon Press Ltd. Printed in Northern Ireland

A theoretical current density ansatz for the quiet day solar semi-diurnal tidal mode of oscillation of the ionosphere*

S. SHANACK†
Lawrence Radiation Laboratory, University of California,
Livermore, California

(*Received* 10 *August* 1959)

Abstract—A theoretical current density ansatz incorporating the atmospheric tidal resonance theory of PEKERIS into the steady-state form of the "generalized Ohm's law", based on a number of simplifying physical assumptions, is obtained. The result is an embodiment of a generalized atmospheric dynamo mechanism of the origin of the upper atmospheric current systems.

The horizontal (co-latitudinal and longitudinal) components of this current density are numerically integrated over an ionized spherical shell concentric with the earth, resulting in approximate expressions for the current density. Various applications of these approximate expressions are mentioned.

1. INTRODUCTION

THE most widely accepted theory of the origin of the upper atmospheric current systems that would produce the quiet day solar and lunar transient terrestrial magnetic variations is the dynamo theory (STEWART, 1882; SCHUSTER, 1889, 1907), in which the atmospheric currents are impelled by electromotive forces induced by the daily tidal periodic motion of the electrically conducting regions across the main geomagnetic field. Equally well known is the mathematical theory of forced tidal oscillations in a rotating atmosphere (PEKERIS, 1937) and its demonstration of the existence of a solar semi-diurnal resonant mode.

It has seemed plausible to attempt to develop, primarily for heuristic purposes, an ionospheric current density ansatz which would incorporate the atmospheric resonance theory into the steady-state expression of the "generalized Ohm's law" (SPITZER, 1956), the latter then being regarded as the embodiment of a generalized atmospheric dynamo mechanism of the origin of the upper atmospheric current systems. In order to develop such an ansatz for the case of the solar semi-diurnal tidal (i.e. gravitational) resonant mode, a number of simplifying assumptions are made.

Firstly, it is assumed that the currents exist only in a thin spherical shell concentric with the earth and located at 100 km above the surface of the earth. Secondly, this shell is supposed to contain equal numbers of electrons and ions per unit volume, that is, the shell is electrically neutral and fully ionized. Thirdly, the regions inside and outside the shell are assumed to consist only of neutral gas atoms. This last assumption has the immediate consequence that the boundary conditions on the inner and outer surfaces of the shell are such that the ion pressure and its radial derivative vanish. The assumption of electrical neutrality within the shell implies that the divergence of the current density vector is zero within the shell.

* Work supported by U. S. Atomic Energy Commission.
† Present address: Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland.

The conductivity in the regions of neutral gas is, of course, zero, while the conductivity and electron density within the shell are taken to be constant.

The above requirement of electrical neutrality within the fully ionized spherical shell leads to a Poisson's equation for the ion pressure P_i, and the vanishing of P_i and the radial derivative of P_i at the inner and outer boundaries of the shell enables the complete solution of P_i to be found in the region where it is assumed to exist. The gradient of this scalar ion pressure P_i then determines the irrotational portion of the generalized steady-state Ohm's law.

2. Formulation of the Poisson's Equation for the Ion Pressure

The following notation is used:

a = radius of the earth,
A = distance from centre of earth to a point 100 km above the surface of the earth;
r = distance from centre of earth;
θ = *co-latitude*;
ϕ = longitude;
g = acceleration of gravity;
ω = angular velocity of the earth;
δ = semi-thickness of the spherical shell;
γ = specific heat ratio;
k = Boltzmann's constant;
m = average mass of atmospheric air molecules;
H = "equivalent depth" corresponding to the solar semi-diurnal resonant mode;
n_e = electron density in spherical shell;
K = specific electrical conductivity in spherical shell;
T = temperature in the isothermal temperature region $(A - \delta) \leqslant r \leqslant (A + \delta)$;
h = scale height in the region $(A - \delta) \leqslant r \leqslant (A + \delta)$;
e/c = e.g.s. e.m.u. charge on the electron;
$\frac{\mu}{A^3}$ = e.g.s. e.m.u. magnitude of magnetic field at 100 km above surface of the earth;
$\rho_0(r)$ = hydrostatic density;
$P(r, \theta, \phi)$ = first-order tidal pressure;
$\Omega(r, \theta, \phi)$ = solar tidal potential;
$u(r, \theta, \phi)$ = southward component of air velocity;
$v(r, \theta, \phi)$ = eastward component of air velocity;
$W(r, \theta, \phi)$ = vertically upward component of air velocity;
$P_i(r, \theta, \phi)$ = ion pressure in the fully ionized spherical shell;
$Q(r, \theta, \phi) = \left(\frac{P}{\rho_0} + \Omega\right)$.

Since the effects of the geomagnetic dipole field upon the total tidal velocity

VOL. 17 1959/6

vector **V** as computed from the linearized hydrodynamical equations are negligible, the components of **V** in a co-ordinate frame fixed relative to the sun may be written:

$$u = 2i \frac{Q(r)}{a\omega} (\cot \theta) Y_2^2$$

$$v = -\frac{Q(r)}{a\omega} (1 + \cos^2 \theta) Y_2^2 \tag{1}$$

$$W = W(r) Y_2^2,$$

where W is identical with the vertical upward component of velocity obtained by PEKERIS. The horizontal components u and v in equation (1) are equal to $\sin^2 \theta$ times the original PEKERIS horizontal components. A slowly varying multiplicative factor, such as $\sin^2 \theta$, was invoked in order to prevent divergences at the poles as a result of the breakdown of the linearization hypothesis. For the case of the solar semi-diurnal resonant tidal mode, $Q(r, \theta, \phi) = Q(r) Y_2^2$, $W(r, \theta, \phi) = W(r) Y_2^2$, where

$$Y_2^2 = \frac{1}{8} \left(\frac{30}{\pi}\right)^{1/2} (\sin^2 \theta) \exp (2i\phi).$$

The "generalized Ohm's law" may be written:

$$\mathbf{J} = K \left[\mathbf{V} \times \mathbf{B}_0 - \frac{c}{en_e} (\rho_0 \nabla \Omega + \nabla \mathbf{P}_i) \right] \tag{2}$$

where $\mathbf{B}_0$ is the geomagnetic dipole field vector, and $\mathbf{J}$ is the current density which by assumption is non zero only within the spherical shell.

Electrical neutrality within the ionized shell implies that:

$$\nabla \cdot \mathbf{J} = 0. \tag{3}$$

In view of equation (3), $\nabla \cdot \mathbf{E} = 0$ also, where $\mathbf{E}$ represents any static electric fields that may exist across the shell. Thus $\mathbf{E}$ does not affect the Poisson's equation obtained by setting the divergence of $\mathbf{J}$ equal to zero, and thus may be considered merely an arbitrary additive vector in the current density $\mathbf{J}$ given by equation (2).

Setting the divergence of $\mathbf{J}$ equal to zero we obtain:

$$\nabla^2 P_i = \frac{en_e}{c} \nabla \cdot (\mathbf{V} \times \mathbf{B}_0) - \nabla \cdot (\rho_0 \nabla \Omega). \tag{4}$$

Expanding the right-hand side of equation (4) by means of the well-known vector differentiation formulae and using the fact that the vector $\mathbf{B}_0$ is irrotational, we have:

$$\nabla^2 P_i = \frac{en_e}{c} \mathbf{B}_0 \cdot \operatorname{curl} \mathbf{V} - \rho_0 \nabla^2 \Omega - \nabla \Omega \cdot \nabla \rho_0. \tag{5}$$

Assuming for simplicity that the earth's magnetic field is due to a dipole at the centre with axis along the geographical axis of the earth, the radially upward, co-latitudinal and longitudinal components of $\mathbf{B}_0$ in the shell are $-(\mu/\pi A^3) \cos \theta$,

$-(\mu/2\pi A^3)\sin\theta$, and zero, respectively. Since also for the tidal mode we are considering $\Omega(r, \theta, \phi) = \Omega(r)Y_2^2$, equation (5) may be written in the form:

$$\nabla^2 P_i = g_1(r)Y_2^2 + g_2(r)Y_2^2\sin^2\theta \tag{6}$$

where $g_1(r)$ and $g_2(r)$ are radial functions given by:

$$g_1(r) = -\frac{en_e}{c\sqrt{\pi}}\frac{\mu}{A^3}\left[\frac{5Q(r)}{a\omega r} + i\frac{W(r)}{r} + \frac{1}{a\omega}\frac{dQ(r)}{dr}\right]\cdot$$
$$-\rho_0(r)\left[\frac{d^2\Omega(r)}{dr^2} + \frac{2}{r}\frac{d\Omega(r)}{dr} - \frac{6\Omega(r)}{r^2}\right] - \frac{d\rho_0(r)}{dr}\frac{d\Omega(r)}{dr} \tag{7a}$$

and

$$g_2(r) = \frac{en_e}{2c\sqrt{\pi}}\frac{\mu}{A^3}\frac{1}{a\omega}\left[\frac{9Q(r)}{r} + \frac{dQ(r)}{dr}\right]. \tag{7b}$$

The particular form of $Q(r)$ and $W(r)$ depends upon the temperature profile in the atmosphere. For a profile consisting of linear parts, in particular, PEKERIS's "Atmosphere B", the explicit expressions for $Q(r)$ and $W(r)$ in the isothermal portion of the temperature profile $T(r)$ for $(A - \delta) \leqslant r \leqslant (A + \delta)$ have been found to be:

$$Q(r) = \frac{i\gamma}{4\omega}gH(1 + V)L\exp\left[-\lambda(r - a)\right] \tag{8}$$

and

$$W(r) = -\gamma HL\left(\frac{1}{2} - \frac{h}{H} + \frac{1}{2}V\right)\exp\left[-\lambda(r - a)\right] - \frac{2i\omega\Omega(r)}{g}. \tag{9}$$

$\Omega(r)$ is taken to be constant $\equiv \Omega$ over the range $a \leqslant r \leqslant (A + \delta)$. The quantities λ, V and L in equations (8) and (9) are constants arising out of the solution of the differential equation for the hydrodynamical divergence and its imposed continuity of solutions for the particular temperature profile in question. For "Atmosphere B", $\lambda = -5{\cdot}69 \times 10^{-2}\,\mathrm{km}^{-1}$, $V = 0{\cdot}273$, $L = i(7{\cdot}22 \times 10^{-8})\,\mathrm{sec}^{-1}$. Further,

$$\rho_0(r) = \rho_0(A - \delta)\exp\left[-\frac{mg}{\gamma k}\frac{1}{T}(r - A + \delta)\right] \tag{10}$$

for $(A - \delta) \leqslant r \leqslant (A + \delta)$. There $\rho_0(A - \delta)$ is the hydrostatic density at an altitude of $(100 - \delta)$ km above the surface of the earth.

Recognizing that $dQ(r)/dr = -\lambda Q(r)$ and that both $d\Omega(r)/dr$ and $d^2\Omega(r)/dr^2$ are negligible in view of the constancy of $\Omega(r)$ over the range $a \leqslant r \leqslant (A + \delta)$, equations (7a) and (7b) become, upon making use of equations (8), (9) and (10):

$$g_1(r) = D_2\left(\frac{D_2}{r} - D_3\right)\exp(-\lambda r) - \frac{D_4}{r} + \frac{D_5}{r^2}\exp\left(-\frac{mg}{\gamma k}\frac{1}{T}r\right) \tag{11a}$$

and

$$g_2(r) = D_6\left(\frac{9}{r} - \lambda\right)\exp(-\lambda r) \tag{11b}$$

VOL. 17 1959/6

where

$$D_1 = -\frac{ien_e}{c\sqrt{\pi}}\frac{\mu}{A^3}\gamma HL \exp(a\lambda)$$

$$D_2 = \frac{5g}{4a\omega^2}(1 + V) - \frac{1}{2} + \frac{h}{H} - \frac{1}{2}V$$

$$D_3 = \frac{g\lambda}{4a\omega^2}(1 + V)$$

$$D_4 = \frac{2en_e}{gc\sqrt{\pi}}\frac{\mu}{A^3}\omega\Omega$$

$$D_5 = 6\Omega\rho_0(A - \delta)\exp\left[\frac{mg}{\gamma k}\frac{1}{T}(A - \delta)\right]$$

$$D_6 = \frac{i\gamma}{8\sqrt{\pi\omega^2}}\frac{en_e}{ac}\frac{\mu}{A^3}gHL(1 + V)\exp(a\lambda).$$

3. Solution of the Ion Pressure Equation

With $g_1(r)$ and $g_2(r)$ given by (11a) and (11b), respectively, the complete solution of equation (6) may be written as:

$$P_i(R, \theta_0, \phi_0) = \iiint_{\tau'} \frac{(g_1 Y_2{}^2 + g_2 Y_2{}^2 \sin^2\theta)\, d\tau'}{|\mathbf{r} - \mathbf{R}|} \tag{12}$$

The surface integrals have vanished in accord with the boundary conditions discussed in the introduction. $\mathbf{R}$ is the radius vector extending from the centre of the earth to the field point, $\mathbf{r}$ the radius vector from the centre of the earth to the "source" point, and $|\mathbf{r} - \mathbf{R}|$ is the magnitude of the vector from the field point to the source point. The ranges of integration for the longitudinal, co-latitudinal, and radial variables are 0 to 2π, 0 to π, and $(A - \delta)$ to $(A + \delta)$, respectively. $d\tau'$ is simply the spherical volume element. The magnitude of the vector from the field point to the source point is given by:

$$|r - R| = \{r^2 + R^2 - 2Rr[\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos(\phi_0 - \phi)]\}^{1/2}$$

where (r, θ, ϕ) are the source point co-ordinates, and (R, θ_0, ϕ_0) are the field point co-ordinates. The real part of $P_i(R, \theta_0, \phi_0)$ is found to be:

$$P_i(R, \theta_0, \phi_0) = \frac{1}{8}\sqrt{\frac{30}{\pi}}\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\, d\theta\, d\phi \left\{\left[(D_1)(D_2 - D_3 r)\right.\right.$$

$$\left.\left. r\exp(-\lambda r) - D_4 r + D_5\exp - \left(\frac{mg}{\gamma k}\frac{1}{T}r\right)\right]\frac{\sin^3\theta\cos 2\phi}{|\mathbf{r} - \mathbf{R}|}\right\}$$

$$+ \frac{1}{8}\sqrt{\frac{(30)}{\pi}}(D_6)\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\, d\theta\, d\phi\left[(9 - \lambda r)\, r\exp(-\lambda r)\right.$$

$$\left.\frac{\sin^5\theta\cos 2\phi}{|\mathbf{r} - \mathbf{R}|}\right] \tag{13}$$

The components of the gradient with respect to the field point of P_i are obtained by straightforward differentiation of equation (13) with respect to the appropriate co-ordinates. The results are somewhat lengthy, and hence will be written down only in the final expression for $\mathbf{J}$, in which $\nabla\mathbf{P}_i$ is readily identifiable.

The components of the first term in equation (2) may be found by employing equation (1) and the explicit components of $\mathbf{B}_0$. The results for the real parts of the radially upward, co-latitudinal and longitudinal components of $\mathbf{V} \times \mathbf{B}_0$ are then:

$$(\mathbf{V} \times \mathbf{B}_0)_r = \frac{1}{8\pi}\sqrt{\left(\frac{15}{32}\right)}\frac{K}{a\omega^2}\frac{\mu}{A^3}\gamma g H|L|(1 + V)\exp\left[-\lambda(r - a)\right] \sin^2\theta\,(1 + \cos^2\theta)\cos 2\phi \quad (14a)$$

$$(\mathbf{V} \times \mathbf{B}_0)_\theta = -\frac{1}{8\pi}\sqrt{\left(\frac{15}{32}\right)}\frac{K}{a\omega^2}\frac{\mu}{A^3}\gamma g H|L|(1 + V)\exp\left[-\lambda(r - a)\right] (1 + \cos^2\theta)\sin 2\theta\cos 2\phi \quad (14b)$$

$$(\mathbf{V} \times \mathbf{B}_0)_\phi = \frac{1}{4\pi}\sqrt{\left(\frac{15}{8}\right)}\frac{\mu}{A^3}\left\{\frac{\gamma g H}{a\omega^2}|L|(1 + V)\exp\left[-\lambda(r - a)\right]\cos^2\theta\sin\theta\sin 2\phi - \left[\gamma H|L|\left(\frac{1}{2} - \frac{h}{H} + \frac{1}{2}V\right)\exp\left[-\lambda(r - a)\right] + \frac{2\omega\Omega}{g}\right]\sin^3\theta\sin 2\phi\right\} \quad (14c)$$

where $|L|$ = magnitude of L.

Finally, the radially upward, co-latitudinal and longitudinal components of $\nabla\Omega(r, \theta, \phi)$ are 0, $(1/r)\sin 2\theta\cos 2\phi$, and $-(2/r)\sin\theta\sin 2\phi$, respectively.

Thus, the components of $\mathbf{J}$ at the field point (R, θ_0, ϕ_0) may be expressed by:

$$\begin{aligned}[\mathbf{J}(R, \theta_0, \phi_0)]_R = {} & \frac{1}{8\pi}\sqrt{\left(\frac{15}{32}\right)}\frac{K}{a\omega^2}\frac{\mu}{A^3}\gamma g H|L|(1 + V)\exp\left[-\lambda(R - a)\right] \\ & \times (1 + \cos^2\theta_0)\sin^2\theta_0\cos 2\phi_0 - \frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\,d\theta\,d\phi \\ & \times \left\{\left[(D_1)(D_2 - D_3 r)r\exp(-\lambda r) - D_4 r + D_5\exp\left(-\frac{mg}{\gamma k}\frac{1}{T}r\right)\right]\right. \\ & \left. \times \frac{r\sin^3\theta\cos 2\phi[\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos(\phi_0 - \phi)] - R\sin^3\theta\cos 2\phi}{|\mathbf{r} - \mathbf{R}|^3}\right\} \\ & - \frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}(D_6)\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\,d\theta\,d\phi\left[(9 - \lambda r)r\exp(-\lambda r)\right. \\ & \left. \times \frac{r\sin^5\theta\cos 2\phi[\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos(\phi_0 - \phi)] - R\sin^5\theta\cos 2\phi}{|\mathbf{r} - \mathbf{R}|^3}\right] \end{aligned} \quad (15a)$$

VOL. 17 1959/6

$$[\mathbf{J}(R, \theta_0, \phi_0)]_{\theta_0} = -\frac{1}{8\pi}\sqrt{\left(\frac{15}{32}\right)}\frac{K}{a\omega^2}\frac{\mu}{A^3}\gamma g H|L|(1+V)\exp\left[-\lambda(R-a)\right]$$

$$\times (1+\cos^2\theta_0)\sin 2\theta_0 \cos 2\phi_0 - \frac{cK}{en_e}\frac{1}{R}\rho_0(A-\delta)\exp\left[-\frac{mg}{\gamma k}\frac{1}{T}(R-A+\delta)\right]$$

$$\times \sin 2\theta_0 \cos 2\phi_0 - \frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi}$$

$$\times dr\, d\theta\, d\phi\left\{r\left[(D_1)(D_2 - D_3 r)r\exp(-\lambda r) - D_4 r + D_5\exp\left(-\frac{mg}{\gamma k}\frac{1}{T}r\right)\right]\right.$$

$$\left.\times \frac{\sin^4\theta\cos\theta_0\cos 2\phi\cos(\phi_0-\phi) - \sin^3\theta\cos\theta\sin\theta_0\cos 2\phi}{|\mathbf{r}-\mathbf{R}|^3}\right\}$$

$$-\frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}(D_6)\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\, d\theta\, d\phi\left[(9-\lambda r)r^2\exp(-\lambda r)\right.$$

$$\left.\times \frac{\sin^6\theta\cos\theta_0\cos 2\phi\cos(\phi_0-\phi) - \sin^5\theta\cos\theta\sin\theta_0\cos 2\phi}{|\mathbf{r}-\mathbf{R}|^3}\right] \tag{15b}$$

$$[\mathbf{J}(R, \theta_0, \phi_0)]_{\phi_0} = \frac{1}{4\pi}\sqrt{\left(\frac{15}{8}\right)}K\frac{\mu}{A^3}\left\{\frac{\gamma g H}{a\omega^2}|L|(1+V)\exp\left[-\lambda(R-a)\right]\cos^2\theta_0\sin\theta_0\right.$$

$$\left.\times \sin 2\phi_0 - \left[\gamma H|L\left(\frac{1}{2} - \frac{h}{H} + \frac{1}{2}V\right)\exp\left[-\lambda(R-a)\right] + \frac{2\omega\Omega}{g}\right]\sin^3\theta_0\sin 2\phi\right\}$$

$$+\frac{2cK}{en_e}\frac{1}{R}\rho_0(A-\delta)\exp\left[-\frac{mg}{\gamma k}\frac{1}{T}(R-A+\delta)\right]\sin\theta_0\sin 2\phi_0$$

$$+\frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}\int_{r=A-\delta}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\, d\theta\, d\phi\left\{r\left[(D_1)(D_2 - D_3 r)r\exp(-\lambda r)\right.\right.$$

$$\left.\left. - D_4 r + D_5\exp\left(-\frac{mg}{\gamma k}\frac{1}{T}r\right)\right]\frac{\sin^4\theta\cos 2\phi\sin(\phi_0-\phi)}{|\mathbf{r}-\mathbf{R}|^3}\right\}$$

$$+\frac{1}{8}\sqrt{\left(\frac{30}{\pi}\right)}\frac{cK}{en_e}(D_6)\int_{r=A-\varrho}^{r=A+\delta}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi} dr\, d\theta\, d\phi$$

$$\times\left[(9-\lambda r)r^2\exp(-\lambda r)\frac{\sin^6\theta\cos 2\phi\sin(\phi_0-\phi)}{|\mathbf{r}-\mathbf{R}|^3}\right]. \tag{15c}$$

4. Discussion and Some Numerical Results

Inspection of equations (15b) and (15c) reveals that the resultant horizontal (that is, co-latitudinal and longitudinal) current density vector vanishes in the northern and southern hemispheres, but not at the equator or the poles.

The angular integrals appearing in (15) can, in principle, be reduced to the Legendre elliptic integrals, but the amount of algebraic manipulation involved is prohibitive. Instead, it is probably quicker to invoke numerical integration over a spherical shell in order to evaluate these triple integrals.

For a shell of semi-thickness of 12·5 km, numerical triple integration yields, at

$R = A$, the following approximate results for $[\mathbf{J}(R, \theta_0, \phi_0)]_{\theta_0}$ and $[\mathbf{J}(R, \theta_0, \phi_0)]_{\phi_0}$ expressed in c.g.s. e.m.u.:

$$[\mathbf{J}(R, \theta_0, \phi_0)]_{\theta_0} \simeq -2{\cdot}84 \times 10^{-12} \sin 2\theta_0 (1 + \cos^2 \theta_0) \cos 2\phi_0$$
$$- 1{\cdot}27 \times 10^{-15} \sin 2\theta_0 \cos 2\phi_0 - 7{\cdot}19 \times 10^{-14} \cos \theta_0 \left[\frac{\cos \phi_0}{(1 - \sin \theta_0 \cos \phi_0)^{3/2}} \right.$$
$$\left. - \frac{\cos \phi_0}{(1 + \sin \theta_0 \cos \phi_0)^{3/2}} - \frac{2 \sin \phi_0}{(1 - \sin \theta_0 \sin \phi_0)^{3/2}} + \frac{2 \sin \phi_0}{(1 + \sin \theta_0 \sin \phi_0)^{3/2}} \right] \quad (16a)$$

and

$$[\mathbf{J}(R, \theta_0, \phi_0)]_{\phi_0} \simeq 1{\cdot}14 \times 10^{-11} \cos^2 \theta_0 \sin \theta_0 \sin 2\phi_0$$
$$+ 5{\cdot}24 \times 10^{-15} \sin^3 \theta_0 \sin 2\phi_0 + 2{\cdot}34 \times 10^{-15} \sin \theta_0 \sin 2\phi_0$$
$$+ 7{\cdot}19 \times 10^{-14} \left[\frac{\sin \phi_0}{(1 - \sin \theta_0 \cos \phi_0)^{3/2}} - \frac{\sin \phi_0}{(1 + \sin \theta_0 \cos \phi_0)^{3/2}} \right.$$
$$\left. + \frac{2 \cos \phi_0}{(1 - \sin \theta_0 \sin \phi_0)^{3/2}} - \frac{2 \cos \phi_0}{(1 + \sin \theta_0 \sin \phi_0)^{3/2}} \right] \quad (16b)$$

5. CONCLUSIONS

From equation (16) the horizontal distribution of **J** based on a generalized atmospheric dynamo mechanism may be plotted onto an ionized spherical shell concentric with the earth at an altitude of 100 km. The upper atmospheric "equivalent current systems" (CHAPMAN and BARTELS, 1940), such as would produce the quiet day solar semi-diurnal transient magnetic variations, are also readily constructed from these theoretical expressions.

Although the physical assumptions from which the above approximate results were derived are no doubt crude, and furthermore that such significant phenomena as thermal tidal effects have been ignored, it is nevertheless clear that the ansatz obtained has a reasonable amount of general structural value associated with it.

REFERENCES

CHAPMAN S. and BARTELS J.	1940	*Geomagnetism* Vol. I and II. Clarendon Press, Oxford.
PEKERIS C. L.	1937	*Proc. Roy. Soc.* A **158,** 650.
SCHUSTER A.	1889	*Phil. Trans. Roy. Soc.* A **180,** 467.
SCHUSTER A.	1907	*Phil. Trans. Roy. Soc.* A **208,** 163.
SPITZER L.	1956	*Physics of Fully Ionized Gases.* Interscience Publishers, New York.
STEWART B.	1882	*Encyc. Brit.* (9th Ed.). 36.

Journal of Atmospheric and Terrestrial Physics, 1960, Vol. 17, pp. 345 to 349. Pergamon Press Ltd. Printed in Northern Ireland

RESEARCH NOTES

Correlation of spread-*F* activity with *F*-region height changes

(*Received* 2 *September* 1959)

Abstract—A new scheme of spread-*F* indices is developed to suit the equatorial spread-*F* observations and applied to estimate the correlation coefficient between spread-*F* activity and *F*-region height changes.

THE phenomenon of spread-*F* is known to be caused by some kind of irregularity, the properties of which are not so far well understood. Some of the recent statistical studies of data taken at different latitudes from the last two decades, made it clear that this phenomenon occurs more frequently at equatorial and polar latitudes.

Spread-*F* at Waltair (17°43′N, 83°18′E) is mainly a night-time phenomenon as is the case with other equatorial stations. In order to make some quantitative studies regular

Table 1. Table of indices

Group height range of the patch (km)	Spread-*F* index
0	0
Up to 25	1
26 to 50	2
51 to 75	3
76 to 100	4
101 to 125	5
126 to 150	6
151 to 175	7
176 to 200	8
201 to 250	9
251 to 300 and above	10

night observations (commencing from 1800 hours IST) were started from November 1958. At each hour the virtual height range over which the echoes from the *F*-region are spreading is obtained at about ten spot frequencies covering most of the *F*-region. In order to avoid the error likely to be caused by the variation of this spread-*F* extent or range with the power changes of the transmitter or the gain settings of the receiver, only echoes having one-tenth of the amplitude of the largest echo in the pattern are considered. For a convenient analysis of the data thus obtained a new system of spread-*F* indices shown in Table 1, is adopted. An index number in the range 0 to 10 is allotted to each of the observations depending on the spread-*F* extent.

Earlier investigators like REBER (1954), BRIGGS (1958) and others adopted a broad classification of the spready h'-f records into three groups depending on the conditions mainly at and near penetration frequencies. But the scheme of indices described above enabled us to follow the hourly, day-to-day, and monthly variations in a greater detail.

This method is convenient at least at equatorial latitudes where on most occasions the entire h'-f record is affected by spread echoes without any speciality near penetration or disappearance frequency; sometimes the trace is affected at all the lower frequencies except near penetration where well defined magneto-ionic components will be observed and f_0F can be determined without any difficulty.

From the data collected so far at this station it is generally observed that the hourly spread-F index obtained as an average of the indices at all the spot frequencies, reaches a maximum value between 2000 and 2100 hours LT. This value of peak spread-F index attained on each individual night is found to be in close correlation with the average rise

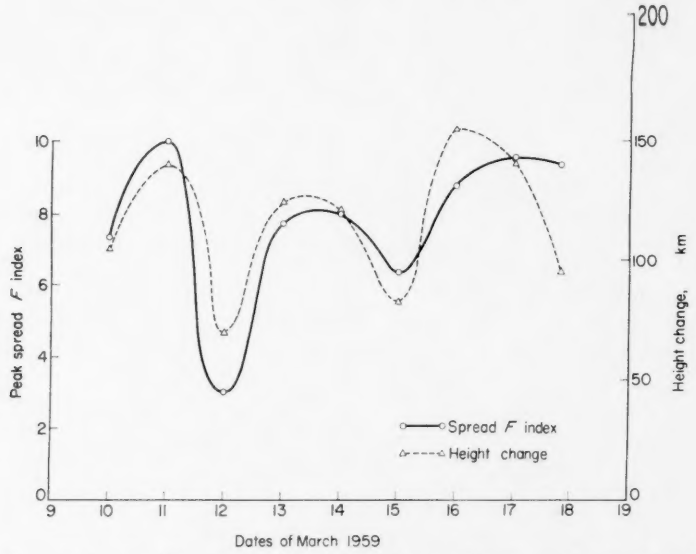


Fig. 1. Day to day variation of peak spread-F index and rise in the F-region height between 1800 hrs and 1900 hrs local time.

in the virtual heights of F-region observed at the low frequency end of the h'-f record between 1800 and 1900 hours. Fig. 1 illustrates this feature showing a close similarity in the day-to-day variations of both the parameters. In order to obtain a quantitative estimate of the degree of correlation about eighty observations were examined. Of these fifty are associated with peak spread-F index greater than zero and the rest with zero. The rise in the lower portion of the F-region was always less than 50 km when the corresponding index is zero. A correlation coefficient of 0·57 is obtained for all the pairs of values considered.

These results appear to be of importance in two ways. Firstly in the prediction of spread-F occurrence and its possible extent on individual nights by a simple examination of the h'-f records taken at 1800 and 1900 hours. These records although completely free from spreading will be able to forecast whether the corresponding night will be free from spread-F occurrence or not. Secondly the results are to be given due consideration in any theory attempting at explaining the origin of the spread-F irregularities. It may be noted that Booker and Wells (1938) reported a similar rise of 100 km or more in the height of the F-region preceding the onset of spread-F at Huancayo.

Further data with detailed discussion of the theories and interpretations will be published shortly.

VOL. 17 1959/6

Acknowledgement—We wish to thank the Council of Scientific and Industrial Research, Government of India, whose equipment was freely used during the course of this work. One of the authors (M. S. V. G.) wishes to acknowledge the council with thanks for his being supported as J.R.A. in one of its schemes.

Ionosphere Research Laboratory,
Andhra University,
Waltair, S. India

M. S. V. Gopala Rao
B. Ramachandra Rao
P. Ramachandra Rao Pant.

References

Booker H. G. and Wells H. W. 1938 *Terr. Magn. Atmos.-elect.* **43**, 249.
Briggs B. H. 1958 *J. Atmosph. Terr. Phys.* **12**, 34.
Reber G. 1954 *J. Geophys. Res.* **59**, 257.

Colour photography of the aurora

(*Received* 22 *September* 1959)

Modern colour films are normally of the subtractive type, with a three-layer emulsion. The spectral sensitivities of these three emulsions may have useful properties for photography of the aurora. For example, the special sensitivities of the three layers making up Kodak "Ektachrome" film are such that simplifying assumptions can be made about the action of light from an aurora on them. Examination of the spectral transmissions of the dyes produced by development in the three layers shows that an image in one of the layers may be viewed separately from the other two images by suitable choice of the colour of the

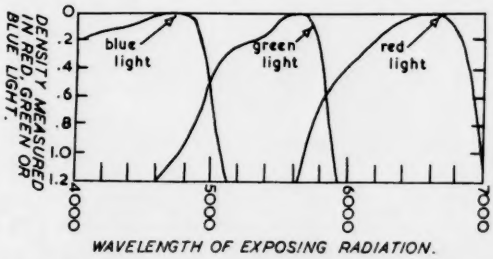


Fig. 1. Photometer measurements of an equal energy spectrum photographed on Kodak Ektachrome E-2 film. The measurements were made using red, green and blue viewing lights, as described in the text.

viewing light. The appropriate bands of wave-lengths used for viewing are, in the case of Ektachrome, 4200–4700 Å, 5100–5500 Å and 6400–6900 Å for the blue-, green- and red-sensitive layers, respectively. Light of wave-lengths in these three regions was used for the photometric measurement of an image of an approximately equal-energy spectrum, the measured optical density being shown in Fig. 1. In plotting this diagram, a small adjustment was made to each curve (by the addition of a constant) to bring the minima of the curves to zero. From these measurements, it is seen that the response of Ektachrome film to the main auroral emission features (the nitrogen bands at 4278 and 3914 Å, and the oxygen lines at 5577 Å and 6300 + 6364 Å) is well represented by considering each layer as sensitive to only one of the emission features. It follows, therefore, that a photograph of an aurora

on Ektachrome film can be considered as three superimposed photographs taken separately through filters isolating the main auroral features. The use of colour film has the following practical advantages over the use of three cameras with isolating filters:

(i) Only one camera lens is required, instead of three, and any imperfections in the lens will cause similar effects in all three images.

(ii) The three photographs are automatically in register.

Characteristic curves for a particular film may be obtained by photographing a wedge spectrum in the usual way and then measuring the optical density of the film with the three

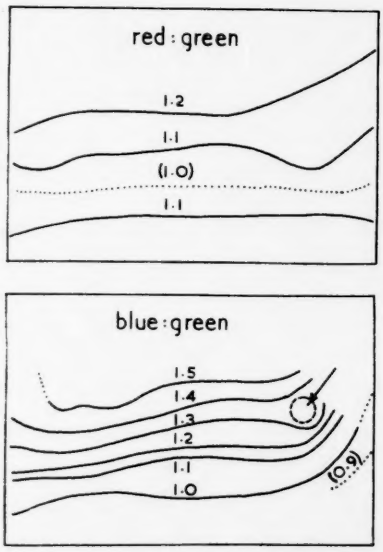


Fig. 2. Contours of relative intensity of auroral emission features in a homogeneous arc. The edges of the diagrams correspond to the edges of the camera gate and are approximately 45° horizontally and 30° vertically.

"viewing" colours. A slightly different method has been used with the colour film measured so far, in that a photometric wedge has been photographed on to the colour film directly. The wedge has been photographed under illumination corresponding exactly in wave-length and, with band systems and the oxygen doublet, in relative intensity to the emission features of the auroral spectrum. The duplication of these emission features was carried out using the Wright spectrophotometer (Wright, 1954) as a template monochromator.

Successful colour photographs of an aurora were obtained on 26 October 1956 at St. Andrews Observatory, Scotland, and measurement of these has provided some interesting results. Using Ektachrome (Process E-2) film and a 45 mm $f/2{\cdot}8$ lens, the homogeneous arc observed early in the display required an exposure of between 17 and 30 min for a satisfactory level of exposure. Because of the uncertainty of possible differences in reciprocity failure of the three layers of the colour film, relative intensities on an arbitrary scale have been measured—that is, the results have been used to determine changes in relative intensity of the main spectral emission features.

Fig. 2 shows the plots of the relative intensities of the red and green, and blue and green emissions in an arc photographed at 1850 hours, UT. The contours of relative intensity are smooth and show a decrease in relative intensity of the 5577 emission with elevation. This is almost certainly the well-known altitude variation of the oxygen green line (Swings,

VOL.
17
1959/6

1952). In the lower diagram of Fig. 2, an area is indicated with an arrow. In this region, the blue : green ratio is higher than would be expected from comparison with the surrounding areas. A similar area was seen on two succeeding photographs taken on nearly the same bearing, and four photographs taken later in the evening show a similar area on each of the four photographs. Fig. 3 shows a plot of the azimuth and elevation of these two regions, measured by reference to the star background, from a total of seven photographs. It is plain from Fig. 3 that the patches are showing a systematic movement from east to west. Since these patches occur in differing positions in the camera gate in the photographs, and since the patch shows up consistently only in the blue–green relative intensity plots, it is

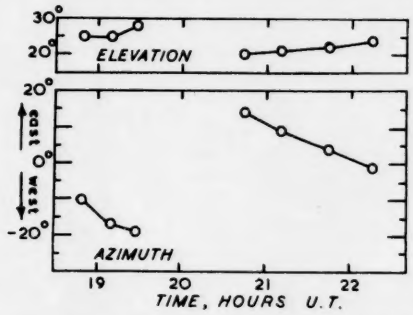


Fig. 3. Plot of the azimuth and elevation angles of the patches of enhanced blue emission.

concluded that they are, in fact, auroral in origin and are not optical or photographic effects or due to a small cloud drifting across the field of view.

The appearance of the patches consistently in the blue–green plots, and not in the red–green, suggests that in these areas the blue emission from the aurora was locally enhanced. This, in turn, suggests that the areas are due to local enhancements of the ionization density, and it is tempting to associate the regions with radar echoing regions. Against this, there are the observations by UNWIN (1959) that the echoing regions are normally extensive in range and azimuth, and the conclusion of GADSDEN (1959) that it is unlikely that there is a close, one-to-one, correspondence between the so-called "radar" aurorae and aurorae observed visually.

In view of the simplicity of the technique of measurement, and since there are now very fast colour films available commercially, this technique may prove to be of value to networks of volunteer observers. The technique was developed at the Imperial College Technical Optics Section, and the encouragement of Professor W. D. WRIGHT is gratefully acknowledged.

M. GADSDEN

D.P.L. Auroral Station,
Awarua Radio, Private Bay,
Invercargill, New Zealand,

REFERENCES

GADSDEN M. 1959 *Ann. Géophys.* **15,** No. 3.
SWINGS P. 1952 *The Atmospheres of the Earth and Planets* (Edited by KUIPER G. P.) p. 187. The University Press, Chicago.
UNWIN R. S. 1959 *Ann. Géophys.* **15,** No. 3.
WRIGHT W. D. 1954 *Optica Acta* **1,** 102.

Papers to be published in future issues

W. C. Hoffman: Conditions for the persistance of purely longitudinal or purely transverse propagation

R. L. Dowden: Ionospheric thermal radiation at radiofrequencies in the auroral zone

G. R. A. Ellis: Gyro splitting of ionospheric echoes

E. D. R. Shearman and J. Harwood: Sporadic-*E* as observed by back-scatter techniques in the United Kingdom

V. Marasigan: Spread-*F* in Baguio through half a solar cycle

T. Obayashi and Y. Hakura: Enhanced ionization in the polar ionosphere associated with geomagnetic storms

R. A. Duncan: The equatorial *F*-region of the ionosphere

J. E. C. Gliddon and P. C. Kendall: Theoretical world curves of maximum ion density in a quiet *F*-region

J. A. Gledhill and A. D. M. Walker: The "Valley effect" in the interpretation of ionospheric eclipse records

R. F. Stebbings, G. V. Groves, R. L. F. Boyd and P. J. Bowen: Low-altitude testing of the rocket-grenade experiment for upper atmosphere temperatures and winds

J. Veldkamp and D. van Sabben: On the current system of solar-flare effects

P. Das Gupta and K. K. Vij: Statistical analysis of fading of a single down-coming wave

D. R. Bates and M. Nicolet: Ion–atom exchange

P. Bandyopadhyay: Electron distribution in the *F*-layer of the ionosphere over Haringhata (Calcutta) on quiet and disturbed days

J. Taubenheim: Incoherency of pulse echoes and the measurement of ionospheric absorption

J. A. Fejer: Hydromagnetic wave propagation in the ionosphere

J. S. Greenhow and J. E. Hall: Diurnal variations of density and scale height in the upper atmosphere

D. Lepechinsky and C. Taieb: La détermination des hauteurs réelles de l'ionisation dans les trés hautes latitudes

G. H. Bazzard: Short-term differences in the behaviour of two daily indices of solar activity during the IGY

J. W. Wright: Some magnetoionic phenomena of the arctic *E*-region

B. J. Robinson: Diurnal variation of the electron distribution in the ionospheric *E*-layer

C. M. Minnis and G. H. Bazzard: A monthly ionospheric index of solar activity based on *F2*-layer ionization at eleven stations

S. R. Khastgir and R. N. Singh: The size of the moving irregularities on the *F*-region and the spread-angle of the radio waves scattered from them

Syun-Ichi Akasofu: On the ionospheric heating by hydromagnetic waves connected with geomagnetic micropulsations

J. S. Greenhow, E. L. Neufeld and C. D. Watkins: The scattering of 36 Mc/s radio waves by weak auroral ionization

VOL.
17
1959/6

OL.
17
9/60